The Foundation of the Concept of Relevance

Birger Hjørland
The Royal School of Library and Information Science, 6 Birketinget, DK-2300 Copenhagen S, Denmark. E-mail: bh@db.dk

In 1975 Tefko Saracevic declared "the subject knowledge view" to be the most fundamental perspective of relevance. This paper examines the assumptions in different views of relevance, including "the system's view" and "the user's view" and offers a reinterpretation of these views. The paper finds that what was regarded as the most fundamental view by Saracevic in 1975 has not since been considered (with very few exceptions). Other views, which are based on less fruitful assumptions, have dominated the discourse on relevance in information retrieval and information science. Many authors have reexamined the concept of relevance in information science, but have neglected the subject knowledge view, hence basic theoretical assumptions seem not to have been properly addressed. It is as urgent now as it was in 1975 seriously to consider "the subject knowledge view" of relevance (which may also be termed "the epistemological view"). The concept of relevance, like other basic concepts, is influenced by overall approaches to information science, such as the cognitive view and the domain-analytic view. There is today a trend toward a social paradigm for information science. This paper offers an understanding of relevance from such a social point of view.

Introduction

In 1975 Tefko Saracevic[1] declared "the subject knowledge view" to be the most fundamental perspective of relevance: "I wish to suggest that the *subject knowledge view of relevance is fundamental to all other views of relevance*, because subject knowledge is fundamental to communication of knowledge. In that lies the importance and urgency of the work on that view" (1975, p. 333; italics in the original). However, since that time almost nobody has repeated this view about the importance of subject knowledge. Also, Saracevic himself seems to have changed his view in his later writings. The few contributions that have taken this approach, such as Hjørland (1997, 2000, 2002; Hjørland & Sejer Christensen, 2002), have so far been met with limited enthusiasm (cf. Saracevic, 2007, p. 1921) and have mainly been overlooked or ignored. We have the strange situation that what is seen as the most fundamental perspective has not been considered further.

If Saracevic was right that "the subject knowledge view" is the most fundamental perspective of relevance, it follows that the competing perspectives on relevance have (implicit) defects—at least I shall argue so. However, these defects have not hindered their diffusion and popularity in information science.

This paper contains arguments that the dominant views on relevance today are less fruitful than the subject knowledge view. I consider this issue related to the developments of metatheoretical discourses such as the cognitive view and the domain-analytic view in information science. The received view is that the user-oriented view of relevance became dominant in the second half of the 1970s. Before that "the systems view of relevance" was the dominant view. I have argued elsewhere (e.g., Hjørland & Albrechtsen, 1995; Hjørland, 2009) that a social paradigm is a better theoretical framework than the cognitive view.[2] This has also been argued by Cronin (2008). This paper develops a view of relevance informed by a social paradigm (termed domain-analysis or the socio-cognitive paradigm), which corresponds with what Saracevic termed "the subject knowledge view." It constitutes the first full-length article attempting to situate "the subject knowledge view" as a foundation for understanding the concept of relevance within information retrieval/information science (IR/IS).

Saracevic (1975) presented and discussed five theoretical frameworks which have influenced thinking about "relevance" in Information Science:

- The system's view of relevance.[3]
- The destination's view of relevance (now commonly termed "the user's view").
- Subject literature view of relevance.
- Subject knowledge view of relevance.
- The pragmatic view of relevance (including "situational relevance").

This article reconsiders each of these views in turn. The pragmatic view will not be discussed under a separate heading but together with "the subject knowledge view" for the reason that I understand it as one view on subject knowledge. Because the main emphasis in the literature is on the dualism

Received July 27, 2009; revised August 31, 2009; accepted September 29, 2009

 DOI: 10.1002/asi.21261

of system's view versus user's view, that dualism will be examined most closely. Documentation for the domination of this dualism in the literature is provided by Saracevic, who wrote:

> In his conclusions, Mizzaro [1997] posits the orientation of works in different periods: "The '1959–1976' period is more oriented toward relevance inherent in documents and query. In the '1977–present' period . . . the researchers try to understand, formalize, and measure a more subjective, dynamic, and multidimensional relevance" (p. 827).
>
> This duality reflects approaches to modeling relevance to this day.
>
> Relevance is a participant in a wider battle royal that started in the 1980s and is still going on. It involves two opposing views or models of IR: systems and users. The user side vehemently criticized the system side. The system side barely noticed that it was attacked (Saracevic, 2007, p. 1924–1925).

It is apparent in the literature that the dichotomy between the system's view versus the user's view is dominating, while the subject knowledge view is neglected. Based on this dichotomy between "system-based relevance" and "user-based relevance," we shall first reconsider these approaches.

"System" or "Algorithmic" Relevance

Saracevic wrote: "The systems viewpoint, obviously, considers IR from the systems' perspective ignoring the user. It is based on a model of IR, called the traditional or laboratory IR model, in which the emphasis is on systems processing information objects and matching them with queries" (Saracevic, 2007, p. 1925), while Borlund (2003, p. 914) wrote: "System or algorithmic relevance, [. . .] describes the relation between the query (terms) and the collection of information objects expressed by the retrieved information object(s)."

It has been a difficult process to describe, analyze, and criticize this view, perhaps because "the system's view" seems to be a term constructed by researchers who consider themselves part of a cognitive, user-oriented view of relevance, as a kind of opposition to their own view.[4] Many of the characteristics that have been used to describe this view have been put forward by the user-oriented view, and many of these seem to be problematic when they are considered from the subject knowledge point of view. This implies that the whole justification of this concept may depend on the interpretation made by the user-oriented view and that we must be very careful in taking over this conceptual framework. We have to make our own re-reading and interpretations of, say, the relevance evaluations made in the Cranfield experiments[5] (which are considered paradigmatic cases for the system's view and have been termed "the archetypal approach," cf. Ellis, 1996, p. 23).

The core issues to be dealt with in this section are:

- What is the meaning of "algorithmic relevance"
- What is meant by "the system"?
- Are algorithms and systems "objective" or "subjective"?
- What are the relations between machine functions and "relevance"?
- Is it correct that, for example, the Cranfield experiments did not consider the users' needs?
- May an alternative interpretation of the Cranfield experiments be that relevance was judged by experts rather than by "users"? (So that the important dichotomy is not a system-user dichotomy but an expert–nonexpert dichotomy?)

The expression "algorithmic relevance" seems to be in need of some clarification.[6] Today, Internet search engines such as Google are based on algorithms. A search engine identifies some documents when a user enters some search terms. In reality, of course, some programmers have designed an algorithm that determines what documents should be given priority. It is thus not the algorithm or the system that determines relevance, but the human programmers behind the engine (whether they intended so or not). In this connection, it is important to consider Saracevic's statement: "The whole point of the evaluation of different algorithms is that they produce different outputs for the same query and from the same set of documents in the system" (Saracevic, 2007, p. 1930).

The system's (i.e., the programmer's) selection is thus not "perfect" or "objective" but is a choice made among many possible choices. It may be more or less clever and more or less suited to different purposes. Therefore, it is subjective (and it is difficult to understand why so many people claim the opposite). Its subjectivity is determined by the programmer's choices, including choices of subject access points, weighting, and utilization of link structures. Also, and perhaps most important: The choice of evaluation methods. Is it based on real users, on experts, or on specific test collections? And so on. In the end, programmers' choices are determined by their knowledge, theoretical views and "paradigms," and different kinds of subjectivity of the relevance of search engines are related to the theoretical principles about their design. In the literature of information science, the bias or "politics" of search engines has been examined and discussed (e.g., by Gerhart, 2004, and by Introna & Nissenbaum, 2000). This issue is an example of the subjectivity of the system. The examination of this kind of subjectivity is only possible from a social perspective, and it is opposed to the claim made by cognitive researchers that "the systems view of relevance" is objective.[7]

A note about the concepts "objective" and "subjective" may be needed at this point. A statement or a representation is subjective if it refers to the opinions, beliefs, and feelings of conviction of this or that individual. It is objective if it is independent of people's opinions, beliefs, and views. Traditionally it is believed that science is objective, and that this objectivity is based on observations ("empirical facts"). Kuhn (1962) and other philosophers, however, have changed that view. They have argued that our perception is dependent on "paradigms" (or views learned in a culture or in a scientific discipline). Knowledge is thus in a way always subjective. This subjectivity is not just individual,

DOI: 10.1002/asi

but often shared by groups of people sharing the same fundamental views. Whether or not a statement may be considered objective cannot be determined by any single observation or procedure (and may always be questioned).

A given method, for example, using journal impact factor (JIF) as an indicator of relevance may be relatively "objective" in the sense that different researchers using this method may produce quite similar results. However, for a given task, JIF may be a less adequate indicator. The decision to use one or another indicator (or set of indicators) should be theoretically justified, which involve the researchers' knowledge and theories and thus also their subjectivity. Many researchers may agree, which imply that they share the same subjective opinion, but that does not guarantee objectivity. In this way even a relatively objective method does not implicate an objective solution.

The term "system" needs to be clarified as well. In the "classical" kind of databases (as known from, for example, Dialog) no such algorithms as in the search engines are at play. When the formulation of the algorithmic view of relevance was made, the classical databases dominated the thinking in information science. Although search algorithms were known, there is no indication that this affected the view of relevance in the Cranfield experiments. In the classical databases the user has to specify the databases, the search terms, the fields to be searched, the Boolean operators, and so on. In a way the users are their own search engine (or programmer). The user can choose between different fields. Some fields (title and abstracts) are normally derived from the documents indexed. Other fields (e.g., descriptors and classification codes) are "value added" elements produced by the database producer (not by the database host, i.e., not by the immediate "information system").[8] By implication it can be said that classical databases put a much greater part of the subjective decisions on the user.

"System" may also refer to a classification system, a thesaurus, or any other kind of controlled vocabulary or knowledge organizing system (KOS), among which the latest and most advanced forms are ontologies. The study and construction of KOS have traditionally been regarded as a core (if not *the* core) of library and information science.[9] Those associated with the user's view of relevance have also neglected that research in knowledge organization may be user-oriented in different ways, and therefore it seems problematic just to consider all the above-mentioned systems as representing "system's view of relevance" as opposed to user's view.

It is therefore strange to speak of *the* systems relevance. First, as we saw, at least three different "systems" are at play:

- the literature being represented,
- the database,[10] and
- the database host.

Each of those "systems" affects the way literature can be selected and retrieved and thus the relevance of the output. The closer we come to full text systems, the more the search possibilities are determined by the documents covered by the databases: The words and concepts used by the authors put definite limits to what can be retrieved by users (or by algorithms). Therefore, it is my claim that the advanced study of information retrieval must be based on the study of the properties of the literature and its representation in subject terminology, which again is represented in databases.[11] This view seems to be in accordance with Saracevic (1975) about the subject literature view of relevance. From a computer science perspective the user faces a computer system. However, from an information science perspective the user faces "the universe of recorded knowledge" or "the information ecology." From this perspective the computer is just a tool used to interact with this universe or ecology. The question of relevance is thus primarily to understand the relation between "user needs" and this entire information ecology, *not* to understand the ergonomic aspects of computer systems.

In the classical databases the relevance of the output is thus influenced by: (1) properties of the documents being represented, (2) decisions including value added services made by the database producer, (3) decisions made by the database host, (4) in the case of links and citations, the citations made to the documents by other documents, and (5) the abilities of the user, including the ability to know what to look for and the ability to utilize the literature and the systems ("information literacy"). Relevance research has traditionally been connected with the design and evaluation of information systems. In order to optimize search strategies as well as to improve the systems, it is important to develop knowledge about the role of each of the five factors mentioned above. It is not just about designing systems, but also about providing users with information about the structure and the content of the information ecology with which they are interacting. This is important to mention because much user-oriented research just asks the user what is relevant without trying to understand these mechanisms behind the output that is being evaluated.

It should also be mentioned that "the system" could be, for example, a public library. Documents are selected, cataloged, classified, and indexed, perhaps influenced by ideas taught at schools of library and information science. Often such ideas are based on the view that the library should be based on an application of objective and neutral principles of management and often ideologies about "user-centered" services are dominant. In the end, however, some priorities have been made, and some information needs are better served than others. Libraries are thus also "subjective" in some ways. Some parts of Library and Information Science (LIS) are about the social roles of libraries, cultural politics, and related issues, and thus closely connected with the view of relevance being defended in the present paper. It is revealing that "relevance research" in LIS has not (yet) considered such perspectives, so cultural policy is about relevance, and vice versa. If managers of a library or of a search engine choose, for example, to make pornography less visible and retrievable, then this might be considered a kind of "cultural policy." Neglecting this seems limiting for the integration of relevance research in the broader field of LIS.

DOI: 10.1002/asi

The relation between machine functions and "relevance" is that machines and algorithms are used, among other things, to select and sort documents and document representations. Different ways and criteria are used in the selecting and sorting processes. Depending on how these processes are performed and on what explicit or implicit criteria are used, the output will differ in content and/or in the way it is sorted and presented. Different ways of selecting and sorting are more or less helpful and relevant to different human tasks. Relevance is thus always connected to human tasks and goals, not to machine or "system" processes. Relevance criteria are needed in order to evaluate machine performance, but are not derived from "systems." They are something we need to bring with us when we consider systems.

How Was (and Is) Relevance Assessed According to "the Systems View"?

Cleverdon (1970) reanalyzed some results from the Cranfield II experiments. The types of search questions discussed were both "realistic" or "real-life questions" and "prepared questions" (which is surprising, given the description of this view from the user-oriented community). Relevance assessments were made by people with different backgrounds, mostly scientists in the field. Each assessor evaluated each document (in full text) on a five-point scale and made qualitative notes about the assessment. Most important is that relevance was evaluated in relation to its possible function for the user[12] because this is directly opposed to how the systems view is mostly being described. The paper further discussed how relevance assessments vary greatly among different assessors. Appendix 1 in Cleverdon (1970) lists the test-questions and the real documents used in the test. This seems important because it makes interpretations of the relevance-assessments possible. This procedure seems different from how it is described by the user-oriented researchers. Borlund, for example, writes: "In this approach [the system-driven] queries are assumed [to be] identical to static information needs, and relevance judgments are assumed to be of an objective nature of whether or not the document is about the query, and can as such be made by any knowledgeable person" (Borlund, 2000, p. 49).

This was in my reading not the assumption made by Cleverdon (1970).[13,14] It therefore seems to be the case that the methodology of the Cranfield experiments—the model standing in opposition to the cognitive view—is too negatively described by researchers from the cognitive school. Cleverdon seems to be fully aware of the variability in relevance assessments. What now seems to be a better dichotomy is that the Cranfield experiments were mainly based on expert evaluations, while the user-based tradition is mainly based on "real users." In the perspective of the user-oriented paradigm, expert assessments are seen as "objective" while "real users'" assessments are seen as subjective. This understanding is problematic. Both experts and "real users" are of course subjective in their judgments (but the experts' judgments are generally much better founded and tend to have much less variation, cf. Saracevic, 2008, p. 776[15]).

Concrete analysis of retrieval failures is a kind of relevance judgment associated with "the systems view." If the task of indexers is to provide good indexing, and if the task of indexing theory is to provide a basis for the development of forms of indexing, then information scientists should be able to evaluate specific errors in indexing. Saracevic wrote:

> A lot can be learned from failure analyses, particularly about human performance. Regrettably, failure tests are no longer conducted, mostly because they are complex, very time consuming, and CANNOT be done by a computer. This type of testing is now relegated to history (Saracevic, 2008, p. 772).

Borlund (2000, pp. 50–52) is also very positive in her description of failure analysis. She expresses some doubts about whether to classify it as "the system driven approach" or "the user oriented approach," but in the end she chooses the first. In my opinion, the basic characteristics of failure analysis are: (1) that it is concrete, it does not just talk about relevance in general, but about specific cases; and (2) that it requires subject knowledge to perform.

The main conclusions from this section are thus:

(1) Systems are either built on a conception and operationalization of relevance or they are based on principles that affect the relevance of the output. Thus, we cannot derive the concept of relevance from systems, but we have to bring it with us when we consider systems.
(2) A given information system is based on (or is a part of) other systems and each subsystem influences the final output, which the user faces. It is important to understand how the output is determined by each element in the information chain. If this is not the case, it is not possible to provide research-based principles about the improvement of the final search result. Relevance is not primarily an issue involving human–computer interaction, but an issue involving human interaction with recorded knowledge as represented in discourses, documents, and languages.
(3) The systems view is often described as "objective," but it has been demonstrated that each element is subjective in one way or another. The issue is not whether they are subjective, but in what way they are subjective—and in what way they should be subjective, i.e., what activities, goals, and interests they are meant to support.
(4) A main criticism from the literature has been that the "system-driven approach treats relevance as a static and objective concept as opposed to the cognitive user-oriented approach that considers relevance to be a subjective individualized mental experience that involves cognitive restructuring" (Borlund, 2003, p. 914). A reading of the original literature provides, however, the impression that researchers in this approach were well aware of the variations in relevance assessments and their dependence on many factors, including subject knowledge and whether full-text or only title were available.
(5) Probably the main criticism of the system-driven approach is that it is not concerned with the users' needs. The

DOI: 10.1002/asi

description above about relevance assessment reported by Cleverdon (1970) shows that this is not the case.
(6) A characteristic not mentioned by the critics of the systems view is that it was often more concrete and more dependent on subject knowledge (and acknowledging the importance of subject knowledge) compared to the user-oriented approach.

Our reading of the research literature has thus provided a picture of what is termed "the system's view" that is different from the received view. Also, the name itself seems to be problematic. We shall now turn toward the cognitive/user-oriented view on relevance.

"Human" (User)-Based Relevance

We have already seen this view described the following way: "In the '1977–present' period [. . .] the researchers try to understand, formalize, and measure a more subjective, dynamic, and multidimensional relevance" (Mizzaro, 1997, p. 827).

In the previous section I questioned the received view of the "systems view" of relevance. This also has implications for analyzing the user-based view. If both views are subjective, then what are the important differences (if any)? It is difficult to present the user-based view[16] of relevance because it seems to be based on assumptions that are not explicitly mentioned or addressed, and therefore we have to expose and discuss these assumptions during our presentation and discussion of this view. Among these assumptions are:

(1) The already presented dualism between systems and users.
(2) A preference for studying "real users" (and a somewhat unclear relation to subject knowledge/ expert views).
(3) A strong tendency to psychologize, e.g., by understanding information needs as inner motivational states, implying that it is the users' satisfaction with the output that should form the basis for designing and evaluating information systems.
(4) An unspoken assumption of universal behavior mechanisms.
(5) That user-based studies are not concrete and that they lack reflexitivity (that the researchers do not address their own information seeking and relevance criteria, but just investigate other people).
(6) The studies have a tendency to study "factors" as unrelated properties of relevance.

There are also some theoretical assumptions that are made explicit in the user-oriented view. Among them are:

(7) That relevance is equal to relevance assessment.
(8) That the user-based view is a subjective view of relevance, and
(9) That relevance goes "beyond topicality" (and the last concept is objective).

All nine points will be addressed below in this order.

(1) We saw above that the literature often defines user-based relevance in opposition to system's relevance. We have already seen that "system" is many different things and that it is problematic to claim that all these systems have not considered the needs of users. There is in contemporary philosophy deep skepticism about subject–object dualism. This is addressed by, for example, pragmatism, hermeneutics, paradigm theory, and much more. Subject and object are rather seen as co-developed, as shaped by language and other cultural elements. Such views have direct importance for understanding users and systems as co-evolved (see, e.g., Winograd & Flores, 1986). It implies a serious criticism of the system/user dichotomy in information science.[17,18]

(2) The user-based view of relevance has a preference for studying "real users," which I shall now address. In order to do so, another dichotomy needs to be introduced and considered: the dichotomy between subject knowledge expert views (sometimes termed "the academy principle") versus the user-based principle. This dichotomy has been much neglected in information science, especially since the influence of "user-based" principles began to dominate the discourses in this field.

The academy principle is known from lexicography, where it means that dictionaries have to be founded on expert knowledge as opposed to the knowledge of lay people or actual use of language. Often, of course, systems may also be designed to be used by experts. In those cases one could say that users and experts are the same. There are, however, even in this case two important reasons to distinguish "users" and "subject knowledge experts." First, even "expert users" may not be the best experts available. There is a strong hierarchy or stratification in science, and very few scientists are considered the leading cognitive authorities in their respective fields. By implication, in designing a local information system for a group of experts, it may be necessary to use internationally recognized authoritative knowledge rather than just relying on the local experts. Second, it makes the principles of information science very confusing if a distinction between "users" and "subject experts" is not maintained. It is important for achieving theoretical clarity to know if information systems should in principle be based on users (in their role as users) or on experts (qua experts).

The academy principle seems to be so evident that it should not be necessary to provide much argumentation. Leading information services such as Medline and leading academic libraries use highly qualified subject specialists. It is in libraries as it is in educational institutions: the higher the level of teaching, the higher the demand for subject knowledge. Lower education and public libraries may to a lesser degree depend on highly trained subject specialists. In order to demonstrate the necessity of the academy principle, the reader just has to consider a specific example, such as, "Is chloroquine a relevant drug in the treatment of malaria?" It is evident that medical and pharmacological knowledge is needed (and much more so if the strengths and weaknesses of many alternatives were to be considered). It is not difficult to see that this example could be generalized to all kinds of

DOI: 10.1002/asi

questions put to information systems. Because knowledge is always updated, knowledge itself changes dynamically, and therefore the dynamic nature of "information needs"[19] and "relevance" is to a very large degree caused by this change in our collective knowledge. In the literature of information science the dynamic nature of "relevance" is, however, often connected to the user, rather than to knowledge itself.

Consider another example. Papers in JASIST are evaluated by subject specialists in information science. The relevance of the papers for the field, for the subscribers and for potential readers, is determined by "peer review," not by psychological or cognitive studies of users (what might that mean?). The "information needs" of users in information science can only be judged on the basis of our collective knowledge, not by studying users as something separate. This is thus again a critique of system/user dualism.

(3) There is a clear connection between the user-based view of relevance and a tendency to psychologize. Saracevic (1975, p. 328) mentions in his discussion of the destination's view that psychology entered information science largely as a result of concerns with relevance judgments. It is thus assumed that the study of psychology, rather than the study of subject knowledge, may reveal something about what information is needed and how information systems should be designed. This is in particular the case in the "cognitive view," the mentalism of which has formerly been criticized by, among others, Frohmann (1990) and Hjørland (1997).

Because the cognitive/user-oriented view is based on mentalism/psychologism, relevance criteria and information needs have mainly been sought in the individual mind. Saracevic, for example, said: "Only the user himself may judge the relevance of the documents to him and his uses, i.e., the judgment of relevance may be subjective" (Saracevic, 1970, pp. 120–122; here cited from Schamber, Eisenberg, & Nilan, 1990, p. 764).

I believe this quote reflects a view that looks specious, but which has done much harm in developing information science. That view is that it is the users' satisfaction which should form the basis for designing and evaluating information systems. A criticism of this view was formulated by Soergel (here cited from Hildreth, 2001):

> Soergel argues that overall improvement in the task performance of the user or resolution of the user's problem is a more appropriate measure of utility than subjective satisfaction with initial search results. Soergel is one of the first researchers to point out that users may be satisfied with less than optimal search results, especially when that assessment is made only at the first moment search results are delivered by the system. More than satisfying the user at this point, "What is needed instead is an attempt to make the user successful" (Soergel,[20] 1976, p. 257).
>
> Soergel's criticisms of satisfaction with search output as a measure of retrieval performance did little to slow or hinder the adoption and embrace of this measure by researchers committed to a user-centered model of system evaluation. Immediately, efforts were made to rescue the concept from Soergel's assault (Hildreth, 2001).

Soergel's view is in line with that of Robert Fugmann, who wrote:

> The user as a non-expert in documentation affairs should not be expected to be able to judge crucial and intricate features of a documentation system such as survival power or recall ratios. If an aircraft producer had the flight properties of a new prototype tested by the passengers as the future users, would one also be inclined to regard such an attitude as an indication of user orientedness? (Fugmann, 1973, p. 363).

Applegate (1993) also realized the problem inherent in the user-based view. He used the term "false positives" about users expressing satisfaction even in cases where the research results are bad! Finally, a recent study by Coiera and Vickland (2008) found that user-provided relevance rankings seem to be of limited to no value when designing a search engine, whereas relevance rankings may have a place in situations in which experts provide rankings.

What seems most critical here is the tendency to consider information needs (or needs in general) and relevance criteria as inner motivational states,[21] which may be regarded as a mistake related to what Gilbert Ryle (1949) termed the "ghost in the machine," a "category-mistake." The expression "the car needs petrol" is of course not an indication that the car has a feeling or an "inner motivational state." The meaning of the word "need" is that the car cannot do what *we* want it to do unless it gets some petrol. There is no reason to believe that the meaning of "need" is different when applied to human beings or to information. We may say that a student needs knowledge about English grammar. Like the car, the student probably has no feeling that he lacks this knowledge, but his teacher may find that his written and spoken English could be improved if he learned principles of English grammar. The teacher may also convince the student about this need and *then* this need may become conscious and an "inner motivational state."

It has important consequences for information science and relevance research whether information needs are understood as inner motivational states or as lack of subject knowledge. In the first case, some kinds of psychological studies are relevant, while in the second case science studies and subject knowledge are relevant. A well-known article in information science is Taylor's (1968) "Question-negotiation and information seeking in libraries," which understands the development of information needs as inner development. Taylor's theory is characterized by a mentalistic approach according to which the "information need" progresses in a relatively independent fashion "inside the head" of the user. It develops continuously and goes through the phases Q1, Q2, Q3, and Q4, from a "visceral information need," via a "conscious need" and a "formalized need" to a "compromised need." Taylor's paper and the underlying view has been criticized by Hjørland (1997, pp. 162–176) and by Nicolaisen (2009). This is in line with a criticism expressed by Bernd Frohmann:

> Mentalism's focus on processes occurring in minds conceals the crucial social context of rules. Since we do not understand the rule we are constructing without understanding

DOI: 10.1002/asi

> its social context, or the way it is embedded in the social world, its point, its purpose, the intentions and interests it serves, in short, the social role of its practice, indexing theory cannot avoid investigation into the historical, economic, political, and social context of the rules in its domain. Mentalism, on the other hand, either erases the social dimension altogether by conceiving rules as operating in disembodied, ahistorical, classless, genderless, and universal minds, or else acknowledges it only by expanding the set of rules of mental processing (Frohman, 1990, p. 96).

The empirical studies of users by adherents of the cognitive/user-based view seem based on the idea that some general patterns can be uncovered (in spite of the strong individual nature of subjective relevance). We shall now consider this idea.

(4) There seems to be a paradox in the user-based view: On the one hand, information needs and relevance criteria are seen as strongly individual and subjective. On the other hand, it is assumed that research about users may uncover some general principles that may be used to design information systems. In other words: There seems to be an unspoken assumption of universal behavior mechanisms that may be uncovered by research done from the cognitive/user-based perspective.

We should expect that an approach which is termed "the user-oriented view" is about how user groups differ (e.g., students, professional practitioners, men and women, scientists, and journalists). However, this is not the case. Research on user groups has to be based on a social rather than on a cognitive theory. It is *not* in research based on the cognitive/user-based view that we find studies of how users differ with respect to relevance criteria.

In the user-based approach there is a tendency to assume that any user may be substituted by any other user. This may be seen by the tendency to choose students as representative for users in general: A group of users is used in an experiment from which the findings are generalized to all users (and all information systems[22]). The tendency to use students as subjects in research instead of "real users" was, for example, a criticism raised by Cleverdon (1971, p. 45) against research done by Saracevic. It is, however, not just the use of students that indicates this hidden universalism. The whole tendency to generalize results from user studies seems to be based on the view and the hope that such studies may provide general principles on how to design information systems. The research questions asked by Schamber et al. (1990, p. 773) seem also to be indicative of such a universalist assumption.[23] The concept of "the user" is in itself an abstraction—corresponding to the abstraction of "the system"—that neglects how users are different (and what specific kind of system they are users of).

Schamber et al. (1990, p. 774) draw three central conclusions about the nature of relevance:

- Relevance is a multidimensional cognitive concept whose meaning is largely dependent on users' perceptions of information and their own information need situations.
- Relevance is a dynamic concept that depends on users' judgments of the quality of the relationship between information and information need at a certain point in time.
- Relevance is a complex but systematic and measurable concept if approached conceptually and operationally from the user's perspective.

These points may thus be seen as an explication of how relevance is basically understood from the cognitive/user-based point of view. I have already problematized the points about users' judgments and users' perspectives by pointing out that users may be satisfied with less than optimal search results: The individual user's view at the time of searching need not be the most relevant view to consider. We shall now consider issues about the dynamic and multidimensional nature of relevance.

First, I have not encountered any publications in the literature claiming that relevance is not dynamic and multidimensional. This is, however, an assertion made by some authors against what they term "the system's view," for example:

> The Cranfield model does not deal with dynamic information needs but treats information needs as a static concept entirely reflected by the user request and search statement. Further, this model uses only binary, topical relevance ignoring the fact that relevance is a multidimensional and dynamic concept (Borlund, 2000, abstract).

Second, some information systems are of course much more dynamic than others, just as some information systems provide much more opportunity for multidimensional search behaviors compared to other systems. A library catalog based on the Dewey Decimal Classification is not a dynamic kind of system, whereas citation databases and folksonomies are very dynamic because new citations in the first case and users in the second case continuously update and change the knowledge organizing system. Bibliographical databases with many kinds of subject access points are more multidimensional than systems such as card catalogs with just a few access points. Characteristics such as "dynamic," "multidimensional," and "interactive" are thus properties provided by the underlying technological development. They are not new important insights provided by the cognitive/user-oriented point of view.

Third, the dynamic and multidimensional nature of relevance is not primarily due to psychological characteristics of individuals (although this is also an important factor). These attributes are primarily due to the nature of knowledge. Knowledge is expanding and changing all the time. New theories replace older theories. Often different theories coexist. Terminologies, concepts, subject-relations, genres, and discourses change dynamically, and as we have seen the problem of relevance is closely related to the individual's interaction with the knowledge ecology. We cannot understand information needs or relevance as developing inside the mind of a user disregarding the development in our collective knowledge. This is, however, what is done by the user-oriented perspective. It represents attempts to uncover a general

DOI: 10.1002/asi

psychological mechanism residing in the mind of each individual user.

(5) Studies of relevance in the user-based tradition are seldom concrete,[24] as was, for example, the formerly presented "failure analysis" from the "systems view," which identified specific errors.[25] They are seldom about real cases and they seldom reflect the information scientist's own experiences. Other disciplines' study of science is often much more concrete. They study what information a given scientist actually used. This is often the case in the philosophy of science, the history of science, and the sociology of science. It should also be considered, however, that when professors or information specialists help students to write theses, such help is concrete. Information specialists are supposed to have (some) concrete knowledge of which information sources are relevant when they help users. However, in research on information needs, such concrete knowledge is seldom in play. It seems as if information scientists just ask what the users regard as relevant, but never discuss whether a given source is relevant or not.[26] They seem to be afraid to reveal their own opinions. They may believe that their research becomes more "objective" this way, but in reality it may just become rather useless and trivial because the researchers cannot interpret users' behavior properly. This criticism can be seen as related to the broader issue about "positivism" versus "interpretivism" in the philosophy of science. The criticism raised here implies that the positivist view should be considered mistaken.

Another problem in relation to the user-oriented view of relevance concerns the principle of reflexivity, that is, the idea that theories in a discipline should apply equally forcefully to the discipline itself and to the individual practitioners of that discipline. This is a principle which has in particular been associated with the so-called "strong program" in the sociology of science and introduced as one of the four basic tenets of that program by Bloor (1976).

Information scientists use relevance criteria when they do research. These relevance criteria should—according to the reflexivity principle—not be understood as different from the relevance criteria used by the users being studied by information scientists.[27] When information scientists do research, they learn something about relevance that should be reflected in their theories, and when they develop theories of relevance, these theories should be reflected in their selection of information.

Howard D. White studied how some prominent information scientists (Marcia J. Bates, Christine L. Borgman, William S. Cooper, Michael H. MacRoberts, Henry Small, Karen Sparck Jones, Don R. Swanson, and Patrick Wilson) recited themselves and others in multiple works over time, thereby providing individual citation profiles. This kind of research is not, as White writes, depersonalized: "One needs domain knowledge to interpret the list of names. If one has that, recitation analysis can be quite engaging—a source of intelligence that, unlike much in information science, is not depersonalized" (White, 2001, p. 87).

It is also worth considering that researchers' citations may be interpreted as indications of what they consider relevant. Each published paper is in reality a kind of case-study in the researchers' relevance criteria. It is just a matter of being able to interpret them. To consider one's own motives to cite other papers—as well as to consider colleagues' way of citing—is to consider relevance criteria in "real life."

(6) The user-based view is also related to the criticism mentioned above of abstracting the categories "system" and "user" to a level in which many different underlying causal mechanisms are mixed together in ways that hinder an adequate understanding.

Table 1 shows 80 relevance factors that have been suggested in the literature (from Schamber, 1994, p. 11). I believe this table demonstrates rather typical aspects of much recent research in relevance done by the cognitive and user-oriented tradition. The table lists 80 factors that have been demonstrated in the literature to have affected relevance judgments. Some factors are rather trivial: If there is disagreement on how to define relevance, of course, the judgment of what is relevant will differ. The same is the case if the evaluation conditions are bad. My main objections are, however: (1) that this research seems to be based on the assumptions that general factors can be identified independently of contexts and that research of this kind can accumulate and at some time in the future provide a general theory of human relevance assessment; and (2) that important and unimportant factors are listed rather mechanically, i.e., without any presentation of theory that may explain their mutual relations and deeper causes. For example, the relevance of a document to a task is mixed up with estimating the relevance of a document given limited information.[28] This issue was well described by Saracevic: "The hypotheses offered [by "the destination's view," i.e., the user's view] in relation to experimentation with relevance judgments by and large concentrated on enumerating and classifying the factors that affect relevance judgments" (Saracevic, 1975, p. 328). We have thus identified a tendency in the user-oriented tradition to study "factors," implying studying them as unrelated properties of relevance rather than as part of systems that determine their interrelatedness.

I believe the subject knowledge view of relevance is able to provide a more satisfactory theoretical framework, which will be considered later.

(7) We have now arrived at a core claim in the user-based relevance view: Equating relevance with relevance judgment. This view is explicitly mentioned by Saracevic (1975, p. 328) among others: "The *destinations view* of relevance has [. . .] equated relevance with relevance judgment" (Saracevic, 1975, p. 329; italics in original).

Closely related is the following quote, which was used as an epigraph by Schamber et al. (1990, p. 755): "There is no such thing as *the* relevance of a document to an information requirement, but rather the relevance judgment of an individual in a specific judging situation recording his judgment [. . .] at a certain point in time" (Rees, 1966, p. 318).

I believe that Rees is right if we interpret his view as being that we should not consider any judgment of relevance as the final truth. We should always be open to revising our

DOI: 10.1002/asi

TABLE 1. Eighty relevance factors suggested in the literature[a] (from Schamber, 1994, p. 11).

Judges	Requests	Documents	Information system	Judgment conditions	Choice of scale
Biases (1)	Diversity of content (1)	Aboutness (3)	Access (item identification) (4)	Breadth of document set (1)	Availability of anchors (1)
Cognitive style (1)	Difficulty level (1)	Accuracy (truth) (3)	Access (subject description) (4)	Definition of relevance (1,2)	Ease of use (1,2)
Concept of relevance (1,2)	Functional ambiguity (1)	Aesthetic value (3)	Access (subject summary) (4)	Order of presentation (1)	Kind of response required (1,2)
Error preference (1)	Specificity/amount of information (1,2)	Authorship (3)	Browsability (4)	Size of document set (1)	Number of rating categories (1)
Expectations regarding distribution (1)	Subject matter (1,2)	Credibility (3)	Comprehensiveness (coverage) (4)	Social pressure toward convergence (1)	Type of scale (1,2)
Formal education (1)	Textual attributes (1)	Difficulty level (1)	Convenience of location (3)	Specification of the task (1,2)	
Intelligence (1)	Weighting of component (3)	Diversity of content (1)	Convenience of hours (3)	Time for judging (1)	
Judging experience (1)		Importance (3)	Cost saving (4)	Use of control judgments (1)	
Judgment attitude (1)		Informativeness (3)	Currency (updating) (4)		
Knowledge/experience (1,2)		Interesting content (3)	Ease of detection of relevance (3)		
Professional involvement (2)		Level of condensation (1,2)	Effort expended (3)		
Research stage (2)		Logical relevance (3)	Flexibility (dynamic interaction) (4)		
Use orientation (1,2)		Novelty (3)	Formatting (scannability) (4)		
Vigilance level (1)		Pertinence (3)	Interfacing (help, orientation) (4)		
		Publication source (3)	Links to external sources (4)		
		Recency (3)	Ordering (subject matter) (4)		
		Scientific “hardness” (1,2)	Physical accessibility (4)		
		Specificity/amount of information (1,2)	Precision of subject output (4)		
		Style (1,3)	Reliability (consistency) (4)		
		Subject matter (1)	Response speed (4)		
		Textual attributes (1)	Selectivity (input choices) (4)		
		Usefulness (2,3)	Simplicity (clarity) (4)		
			Time spent (3).		

[a]Based on (1) Cuadra & Katter (1967a), (2) Rees & Schultz [1967], (3) Cooper (1971, 1973), and (4) Taylor (1986). Includes all factors suggested by Cuadra & Katter and selected factors suggested by others.

view of what is relevant in a given case. We should be open to the principle of fallibilism and to the view that different competing views exist in any domain most of the time. When that is said, I believe that the two quotes are based on a very problematic understanding. A basic purpose of research and higher education is to put educated people in a better position to solve problems, and this is closely related to the ability to evaluate the relevance of different alternatives. We are back to the academy principle. The claim that there are just different views of relevance (implying that we should only consider the views of an actual user before us, or a sample of actual users or potential users) should not be accepted. This is obvious if we take a concrete example, for example, from medicine: The user normally does not have the qualifications to evaluate his or her information need.[29]

We may thus conclude that a core view in "the user's view of relevance," that relevance should be equated with relevance judgment, is problematic.[30] It should be recognized that information systems are made to support users. They should be "user friendly." However, user friendliness should not be mixed up with the view that the relevance of some information for a given task should be established by studying users. There seems to be a paradox here: Information scientists are assumed to be professional concerning information systems, services, and sources. They are professionals, while users (qua users) are amateurs or novices. Why should professionals base their professional knowledge on that of novices?

(8) We have seen that perhaps the most important claim about the difference between "the system's view" of relevance and "the user's view" of relevance is that the first is "objective" while the second is "subjective," implying that one of the most important characteristics of the user-based view of relevance is that it is a subjective view of relevance. What is meant by this? And what are the implications of this view?

One simplistic interpretation could be somewhat like this: In the literature experts publish the objective truth. All available documents are thus true and objective, and the user should not care about this. Among the many versions, some documents might be preferred by the user. In other words: a range of qualities, which do not concern the truth, may affect the users' choice of documents. Put in another way: The important aspects of documents (how their knowledge claims are substantiated) are outside the concern of the user as well as outside the tasks of information science. The less important qualities are what information science and users should be concerned with. Given the truth and objectivity of both content and "topicality" information science should investigate users' subjective preferences, to put it in an extreme way: Whether users prefer, for example, to receive print on white paper or on green paper.

I believe the view that this simplistic interpretation uncovers is totally wrong and extremely dangerous. Modern philosophy of science (post-Kuhn) does not see science as the objective accumulation of truth, but as more or less influenced by theories, interests, and cultural issues. It is important that information systems provide information that relates the single claim to its theoretical premises. Also, terminology, concepts, and subject assignment (aboutness, topicality) are related to "paradigms." This view also has implications for the choice of the less important things mentioned in the simplistic interpretation. Other, less important qualities are not randomly distributed. They are related to the core qualities of documents. Designing information systems based on users' preferences is not just an innocent act, but may provide the user with false information.

Consider, for example, Al Gore's film *An Inconvenient Truth* (2006) in which two studies were cited in order to explain why so many people remain skeptical about global warming. The first study looked at a random sample of almost 1,000 abstracts on climate change in peer-reviewed scientific journals from 1993 to 2003 and found that exactly zero doubted "that we're causing global warming." The second surveyed a random sample of more than 600 articles about global warming in popular media between 1988 and 2002 and discovered that 53% questioned "that we're causing global warming." This film is of course not a scientific documentation and its conclusion is simplified compared with the literature in this field.[31] It is, however, sufficient to demonstrate the danger of believing that information may just be selected from the most "user friendly" sources.

If users do not know this, they might perhaps prefer popular media.[32] They may believe that there is no systematic connection between "truth" and other document characteristics, but Al Gore's example shows that this is wrong. Given such knowledge, the implication must be that it is dangerous if information systems are constructed according to empirical studies of users.

(9) In the user-based view of relevance the view is often expressed that relevance goes "beyond topicality."[33] In addition, topicality is seen as a "system-based" and an "objective" phenomenon.

Schamber et al. wrote:

> 2.2. System-oriented relevance: Topicality
> In the sense that relevance is concerned with whether the topic of a query matches the topic of a document, the clearest and most widely accepted definition of relevance has been one of topicality, or the "best match" principle. From the beginning, information scientists have acted on the premise that information systems can be designed, through careful attention to relationships among subject terms in indexing and search operations, to retrieve documents that are relevant to users. At the same time, they have also recognized that topicality is an incomplete definition of relevance. Why does the system-oriented view of topicality relevance persist and, given its persistence, how valuable is it? (Schamber et al., 1990, p. 758).

This is (again) an example of how the literature mixes together a number of assumptions without analyzing them separately, which makes it hard to deal with. "Best match" is introduced, but it is not argued why the understanding of "topicality" is connected to best match (rather than exact match or just the users' understanding of terms in printed

DOI: 10.1002/asi

bibliographies). The concept of "the subject of a document" or "the aboutness of a document" or "the topic of a document" is not, in my opinion, dependent on whether the document is in your hand, whether it is represented in a bibliographical database, or whether it is retrieved by using Dewey Decimal Classification or by free text searching, by best match algorithms, etc. The appropriateness of, say, the title of a paper may be addressed both in general and in relation to retrieval systems. There may be different opinions about whether a given title is appropriate or not, whether or not it is a good description of what the book is about, its subject or "topic." But what has this to do with the "system-oriented" view?

Schamber et al. (1990, p. 758) claim: "Today it is generally agreed that direct topical matches can be measured." I believe that the reasoning behind this quote is wrong and simply neglects a large body of literature claiming the opposite. As authority for this claim they provide just two references, one written by Eisenberg himself and one in press at the time. I believe it is a basic semiotic phenomenon that humans understand words differently and that this is generally accepted in much wider circles than among information scientists. There are many studies of inter-indexer consistency of relevance judgments and subject assignment to documents in LIS demonstrating the opposite of this claim. I shall not provide a review of such studies here, but shall limit myself to referring to Saracevic (2008).

In LIS there has been much research related to "subject," "topic," "aboutness," and related terms. It is a major body of literature, which obviously is not considered in the two quotes. The most important thing to say is that these concepts have been much debated and that some researchers suggest objective understandings of them, while others (including myself) suggest subjective understandings. The assignment of subject categories will always be a subjective process. It may be believed to be "objective," but in reality it must reflect the view of an indexer or programmer or who else is doing this. The subjective nature of topicality is most clear in so-called "request oriented indexing," which is indexing in which the anticipated requests from users influence how documents are indexed. The indexer asks: "Under which descriptors should this entity be found?" and "Think of all the possible queries and decide for which ones the entity at hand is relevant" (Soergel, 1985, p. 230).

I shall not consider this issue in any detail, but conclude; (1) that "topicality" cannot be considered objective, and (2) that there seems to be no reason to connect "topicality" with the system's view of relevance. A contribution by Kari Holland was helpful for my understanding of how people from the user's view may think of this. In response to an example provided by Hjørland (a patient suffers a well-defined disease such as scurvy caused by lack of vitamin C, the relevant medical treatment for him would be doses of tablets containing vitamin C/ascorbic acid), Holland wrote:

> There may however be factors that are not a part of an objective view of relevance, but which are equally important to the user and we need to include both in discussions, and research, on relevance.
>
> For instance, also with scientifically proven knowledge, there are situations where there is more than one solution to a problem. Scurvy can be cured by C-vitamin pills, but it is also cured by a diet of fresh meat, for instance seal meat. If the scurvy patient is in Northern Greenland, one might choose seal meat and not C-vitamin pills. Both are equally relevant according to scientific tests, but situational factors might determine which solution to choose, and it will be the individuals in the situation that make the ultimate decision. And this applies likewise to information retrieval (Holland, 2005).

As an answer to Holland's argument, I will say that knowledge production in society is organized in such a way that chemists isolate and describe chemical materials and their distribution in the environment, while biologists, pharmacologists, and physicians examine causes and cures for illnesses, i.e., chemicals' relations to health and illness (the relation is seldom as clear as in vitamin C and scurvy). Hence, seeking the causes and treatments of scurvy must typically be done by means of some information sources, while seeking information about available sources of vitamin C must be found in other sources. Some sources may provide the combined information (and thus seem more user-oriented). However, such a combination of sources will always be secondary and involve a number of backdrops. It is better for information-literate persons (including intermediaries) to know about how knowledge production is organized and utilize the best available sources and then make the combination. At least it should be clear that information systems for professionals should not make combinations of this kind. We are back to the problem of which information systems we are talking about.

I shall not go further into this issue here, but just summarize some basic aspects. (1) Topicality (aboutness or subject matter) is not something a document "has," but is something ascribed to documents by persons with some qualifications from some perspective and for some purpose. (2) Topicality is thus "subjective" rather than objective.[34] (3) Topicality cannot be said to belong to the systems view (or to any other view for that matter). Different principles and techniques (including algorithms) will influence the subject assignment, and such different assignments may be more or less useful for different purposes. (4) The way topicality has been understood as objective and belonging to the systems view is harmful, because by implication this is considered unproblematic, while improvements may concentrate on less important aspects such as availability, price, novelty, etc.

The following is a summary of our characteristics of the user-based/cognitive view on relevance:

- It establishes a dichotomy between "system" and "user."
- The concept of "system" is very abstract, general, and not clearly described. No clear discrimination between the different "systems" at play has been suggested or discussed. (Such as the literature being represented, the database, the knowledge organizing system, and the database host).

DOI: 10.1002/asi

- The concept of "user" is also very abstract and general. Users are not categorized according to different kinds of expertise or information needs and the need of different relevance criteria. There is no focus on how user groups differ.
- The paradox that, on the one hand, a user's relevance is considered purely individual, while on the other hand the researchers try to uncover general tendencies.
- It focuses on "ordinary users" (not experts) without providing theoretical reasons for this.
- It is associated with a cognitive/psychological view of information needs rather than with a social view or a view related to scholarly methodology or theories of knowledge.
- It is connected to the assumption that any user may be substituted by any other user (an unspoken assumption of universal behavior mechanisms).
- Its interest in users may be interpreted as related to commercial-like interest in customers (and New Public Management philosophy, cf. Kann-Christensen & Andersen, 2009).
- It has a tendency to focus on human-computer interactions, rather than on humans' interactions with the accumulated knowledge, the literature, or the knowledge ecology.
- Studies are not concrete and lack reflexivity. They seem to be based on norms of neutrality and objectivity.
- Based on positivist rather than on interpretative approaches. (Failures, for example, are not interpreted or discussed).
- Rather atheoretical. No attempt to uncover underlying causal relations.
- Antirealist: Users' assumptions are not interpreted in relation to other views or "cognitive authorities." Users' assumptions are identical with what is relevant.
- While subjectivity of users' relevance criteria is recognized, the subjectivity of "topicality" is not. The implication may be that focus is taken away from core issues in information delivery so that only more superficial issues are considered.

I have now finished the critical examination of "the received view" of relevance and am turning to what I consider more fruitful alternatives.

Subject Literature View of Relevance

Saracevic (1975) describes subject literatures as governed by mechanisms of a Darwinian nature: success breeds success. Relevance is what is beneath bibliometric distributions and the structure of the subject literature itself is determined by relevance. The subject literature view of relevance is important, he writes: "For information systems, the process starts with subject literatures. The aim of information systems is to enable and to enhance contact between subject literature and user-destinations" (Saracevic, 1975, p. 331).

I believe this is an extremely important observation that has been almost totally forgotten since "the user's view" and/or "the cognitive view" became dominant. This principle may not cover all forms of "information systems," but it certainly covers bibliographical databases and full-text databases, which are the kinds of databases that have totally dominated research in Information Science and Information Retrieval.

What should be considered relevant means: To be considered relevant in scholarly arguments. Scholarly domains—and scholarship in general—develop criteria of what counts as serious evidence and arguments. It is widely accepted that such evidence must be published in order to be properly documented. The nature of scholarly documentation puts demands on scholarly communication, terminology, concepts, and the publication system—as Saracevic realized. "Information systems" (i.e., bibliographical databases) are elements in the same scholarly systems, and what is considered "relevant" in the databases must be based on what is considered relevant in the primary information system (i.e., journals and monographs).

I believe there has been much dislike of this view in IS. Scholarship is seen as too narrow compared with all the purposes for which "information systems" may be applied. Lack of motivation to engage in scholarly fields may be another reason.[35] However, I believe that the problems in information retrieval are such as to make Saracevic's 1975 insight decisive, and that lack of progress is caused by the neglect of these important insights. The subject literature view of relevance is in agreement with "the subject knowledge view" (and almost identical with it[36]). This is why we shall now turn to that view.

Subject Knowledge View of Relevance

Saracevic (1975, p. 331f.) mentioned a distinction between "relevance" and "pertinence" suggested in the early 1960s and also used by Foskett (1972). In Foskett's presentation, "relevance" refers to the consensus or paradigm in a given field, while "pertinence" refers to the individual user's need. A system may provide document representations which are relevant, but if the user already knows them they are not pertinent in relation to her/him. It is evident that this terminology did not catch on. The user-oriented view of relevance does not generally use "pertinence" for its view, and the consensus view (or subject knowledge view)—as formerly stated—has been almost totally neglected. I shall not consider the term "pertinence" any more, but just speak of "relevance" in both cases.

Foskett's view of relevance is close to the one suggested in the present article. Before I did so, he considered the literature of science studies and found here the basis for understanding "relevance": That relevance is not primarily a psychological concept, but a concept in the theory of knowledge (epistemology), a paradigm, something generally accepted in a community. Consider the following quote from Cleverdon, Mills, & Keen (1966) about what went wrong in "the ASTIA-Uniterm test":

> Up to this point, everything appears to have gone according to plan. The final stage was intended to be a comparison of the output of the two sets of searches, in order to find which system had been successful in obtaining more relevant documents.
>
> The problem which arose at this final stage was that neither group was willing to accept the relevance assessments of the

DOI: 10.1002/asi

other group; rumour has it that at the end of the second day of discussion, the two groups were still arguing about the meaning of the first search question. No real blame can be fixed on those who organized the test; in 1952 it was not unreasonable to think that two groups of intelligent people would, without serious difficulty, be able to come to an amicable agreement as to which documents were relevant to a particular question. If any fault can be found, it only lies in the failure to make generally available either of the two reports which are said to have been prepared by the two groups taking part in the test (vol. 1, part 1, pp. 9–10).

This quote illustrates what seems reasonable in the "post-Kuhnian" period: That different views or "paradigms" tend to exist in any domain, and that disagreements and subjectivity should not just be understood as purely individual, but can be of a more systematic nature related to basic assumptions shared by a number of people (which may even be organized according to such views). Cleverdon et al. are fully correct in asking for the reports. Such dissenting reports may be interpreted by subsequent researchers and thus provide better understanding of underlying relevance disagreements. The reason this have not be done may be (1) that such kinds of interpretation are not seen as proper "scientific" by positivist norms, and (2) that it requires subject knowledge and therefore is difficult to do. The approach in the "subject knowledge view" is thus to search for variations in relevance assessments that are connected to basic views or "paradigms" in a domain or across domains. This opens a third way between universal relevance assumptions, at one end, and purely individualistic relevance assumptions at the other end.

Relevance in this article is understood in relation to human activities: Something is relevant in relation to a task if it supports the fulfillment of that task.[37] Relevance in LIS is about what kinds of knowledge or information are fruitful for different purposes: What kinds of knowledge serve as a tool for some human actions?

In society many things, such as educational programs, are based on implicit or explicit views of relevance: When we grew up, we were socialized into established norms about relevance. We chose, for example, an education based on a view of its relevance for our future job opportunities. Some norms seem very well based, others seem less so. Some kinds of knowledge have more or less recognized importance for solving specific tasks; for example, mathematics has such a recognized status in physics. This is an example of recognized relevance, although that also may change dynamically over time. Other fields such as the social sciences have much less consensus regarding the kind of information relevant to their further development. Theoretically the relevance of some knowledge to a given task is never "given," but is more or less well established and dependent on theoretical assumptions.

What is seen as relevant in a field is very much dependent on the theoretical outlook: If, for example, you believe schizophrenia is caused by double-bind communication in childhood, then family studies—and the information they can provide—become relevant. If, on the other hand, you believe that schizophrenia is caused by chromosome disorders, then genetic studies become relevant. To a large degree research is organized according to theoretical outlooks determined by such things as disciplinary borders, preferences for specific research methods, "paradigms" etc. This does not imply total relativism: Relevance is dependent on theoretical outlook, but theories cannot be chosen randomly but have to be examined and justified.

The question of seeking "relevant information" is both a question of (1) developing or choosing a fruitful theoretical approach to the task in question, and (2) of knowing how the relevant information is distributed over disciplines, conceptions, publications, etc. (i.e., "the information ecology"). (1) and (2) cannot be learned independently, but are mutually related: When you discovered that family communication probably is a bad theory for explaining schizophrenia and changed your belief to the genetic theory, then you also had to change your ideas about where to find information (new disciplines, new concepts, new databases, new citation networks, new genres, etc.[38]).

Fundamentally a given person's view of what is relevant for a specific task is thus determined by his or her theoretical outlook. It is "social" in several ways:

(1) The concepts and knowledge used by an individual are learned in a social setting and influence his or her outlook.
(2) Different views compete in communities and are often associated with different social interests and ideologies.
(3) Individuals may depend on authorities rather than form independent opinions.[39]

In the perspective of pragmatic philosophy[40] subject knowledge is seen as something that is constructed in order to serve some human goals. Medical knowledge, for example, helps cure people of diseases. Information science knowledge helps people find relevant information, etc. The relevance of knowledge depends on its usefulness to achieve specific goals. Based on this understanding, I provided the following definition of relevance:

> Something (A) is relevant to a task (T) if it increases the likelihood of accomplishing the goal (G) which is implied by T. (Hjørland & Sejer Christensen, 2002).

This definition applies to anything, including documents. As such it is clear cut.[41] What in practice causes problems are theoretical disagreements on what the goals are, on what criteria for good solutions are, and on what methods are available. Goals and problems may be differently conceptualized and connected to different world views and epistemologies.

Consider somebody handing a document to, for example, a historian or a medical researcher, who looks at the knowledge claims in the document and how they are substantiated. The professional reader quickly considers:[42]

(1) What are the theoretical assumptions on which this document is based? (Do they seem to be well informed or naïve? How do they fit with current thoughts—and with my own ideas?) If the paper seems to be based on new

DOI: 10.1002/asi

ideas matching those of the user, then the paper will normally be considered relevant (given that the next question can also be answered positively[43]).

(2) Are the claims in the document based on proper scholarly arguments and research methods? If they are based on what that professional considers poor research methods, the document is normally considered "not relevant."

Some readers may find that this example is too narrow: that relevance is much broader than scholarly arguments. They may believe that ordinary people use other criteria and that the academic criteria are much too narrow to be important in the larger picture. That may be the reason why information scientists have been much more concerned with psychology and "users" than with science studies and cognitive authorities. I believe this view is fundamentally wrong: People in general act according to how they see the world and what they trust. The first is about theories of reality, the last is about theories of knowledge. When we grow up in a culture, we acquire—both implicitly and explicitly—theories of both kinds.

Think, for example, about the role of women in society. One older view was that women were unfit for many jobs (but fit for housework) and that women should not be allowed to vote, etc. This view was not just challenged by academic feminists, but by all the women who, for example, entered science, management, and other jobs considered for men. In doing so they challenged the established view of gender roles, and they found the kind of theoretical support used by the old view to be nonrelevant. These pioneers challenged the established views on what kind of knowledge is relevant about female psychology. Also, those women, who did not do research, were thus involved in a paradigmatic struggle between different views of what women can do and what they should do.

In all scholarly fields students are taught concepts, theories, and research methods. The methods are not just relevant for doing research: Methods are equally relevant for reading research. They are relevant for interpreting the research results (i.e., evaluating the knowledge claims in the papers). In other words, professionals are taught how to determine which claims seem well substantiated and therefore relevant. There are always too many hypotheses to go along with, and therefore many have to be dismissed. All people have to consider a tradeoff between openness to new views, on the one hand, and, on the other hand, to seriously consider the implications of some views (and thus be restricted rather than open-minded). People have to limit the number of hypotheses they consider fruitful, and relevance is largely determined by the believability of those hypotheses. Relevancy is narrowly connected to which hypotheses have a level of trustworthiness that makes them worth further consideration. This has much to do with cognitive authority and with scholarliness.

Today's view of knowledge and of the scholarly method is no longer the same as during the heyday of logical positivism. The epistemology of logical positivism was empiricist and rationalist, neglecting the theoretical, cultural, conceptual, and pragmatic aspects of the knowledge. That changed in what we may term "the post-Kuhnian" period (i.e., after Kuhn, 1962). This implies that there are now more ways to teach research methods and more ways to read papers.[44] An interpretation of the status of the theory of knowledge (and thus relevance assessment) can be formulated in a few points:

- Logical positivism tried in the first half of the twentieth century to develop general and objective research methods which at the same time implied criteria for what to consider relevant knowledge claims (all knowledge was considered scientific).
- Post-Kuhnian epistemology has given up many of the ideals of logical positivism[45] and thus provided less general and less clear-cut methods and also made relevance assessment less clear-cut (and made people's relevance criteria more divided).
- Research methodology, science studies, and epistemology have developed important insights. Such insights are the proper basis for a theory about judging relevance. This kind of insight is not in "the head of the users" but first and foremost embedded in scientific practice, in the scientific literature, and in the metatheoretical and the critical literature.
- Epistemology is also about the relevance of "alternative medicine," "high" versus "low" culture, feminist criticism of science and of everyday life, the political ideologies of scientists, the relation between science and society, etc. All such questions form the basis of "relevance" (see Abrahamsen, 2003, and Ørom, 2003, as fine examples from LIS).
- Theory of knowledge (epistemology) is thus the core subject of relevance research in information science.

There are many theories of knowledge at play today (see Hjørland & Nicolaisen, 2005, for an overview). Many views are related to each other and thus form "families" of related views. In order to summarize theories of knowledge and their associated views of relevance, Hjørland (2002) provided Table 2.

How Does the "Subject Knowledge View" Relate to the IR Situation?

When, for example, historians are looking for information about a specific event, they use knowledge about source criticism developed within the field of history (see Hjørland, 2008b). The historian tends, for example, to prefer primary sources to secondary sources, to use information based on relics rather than information based on narratives, etc. If the historian is a feminist scholar, s/he may prefer sources and interpretations reflecting a feminist-epistemological understanding.[46] When searching databases, they search for clues (or "signs") that are likely to reveal something about the trustworthiness of the documents, including methodological terms, esteemed authors and publishers, and much more. "The ordinary users," on the other hand, are less aware of such methodological issues and therefore tend to select documents from more superficial characteristics or clues, e.g., the newest ones, the more popular ones, the cheapest ones, etc. It is thus easy to understand why experts' judgments generally tend to have much less variation (cf. Saracevic, 2008, p. 776). It is, however, important to consider that we should expect

DOI: 10.1002/asi

TABLE 2. Simplified relevance criteria in four epistemological schools (from Hjørland, 2002, p. 269).

Empiricism	Rationalism	Historicism	Pragmatism
Relevant: Observations, sense-data. Induction from collections of observational data. Intersubjectively controlled data.	Relevant: Pure thinking, logic, mathematical models, computer modeling, systems of axioms, definitions and theorems.	Relevant: Background knowledge about pre-understanding, theories, conceptions, contexts, historical developments and evolutionary perspectives.	Relevant: Information about goals and values and consequences both involving the researcher and the object of research (subject and object).
Nonrelevant: Speculations, knowledge transmitted from authorities. "Book knowledge" ("reading nature, not books"). Data about the observers' assumptions and pre-understanding.	Low priority is given to empirical data because such data must be organized in accordance with principles which cannot come from experience.	Low priority is given to decontextualized data of which the meanings cannot be interpreted. Intersubjectively controlled data are often seen as trivia.	Low priority (or outright suspicion) is given to claimed value free or neutral information. For example, feminist epistemology is suspicious about the neutrality of information produced in a male-dominated society.

different degrees of agreement and clusters of agreement among people sharing the same epistemological theories. To be inspired by, say, "postmodernism," means—more or less—to share a view of what is considered relevant.[47] To be confused about epistemology means to be confused about relevance.

Epistemology is about the fundamental issue in relevance assessments. Of course many other criteria are involved, but not with the same importance. A theory of relevance must determine what essential issues are and what superficial issues are. Purely individual/idiosyncratic views of relevance are not of much use as guidelines for designing information systems and services.

Conclusion

The "received view" of relevance in information science is based on a system–user dualism, against which arguments have been put forward in this paper. The distinction between system's relevance and user's relevance is considered defunct because relevance is only meaningful in relation to goals and tasks, and machines do not have goals. Algorithms and systems may select or sort items in ways that are more or less adequate for supporting different human activities, i.e., that are more or less relevant in relation to given tasks. Relevance is thus never "a system's," but always "human" and therefore the dichotomy is wrong. To determine which items are relevant in relation to a given goal/task requires subject knowledge and is dependent on different theories/views. Users of information systems are therefore not automatically competent to judge relevance.

The subject knowledge view represents an alternative to the received dichotomy. This view can be related to domain analysis in which systems and users are seen as co-developed and influenced by the same theories. Consider a feminist library (i.e., "a system") and a feminist scholar (i.e., "a user"). They both share some common concepts, terminology, and ways of seeing the world. They also share relevance criteria. When a feminist scholar searches a feminist database, it is not an abstract and isolated "user" confronting an independent and abstract "system." It is unfruitful to approach relevance from this dualist perspective. The key to understanding how to develop the feminist database is not by studying some arbitrary users, not even feminist scholars qua users. The key to the development of the feminist database is in the conceptions and theories developed in feminist scholarship: A collective knowledge domain.

While Saracevic (1975) discussed pragmatism as one of five overall views of relevance, he did not discuss alternative views of knowledge. He did not ask: If pragmatism is one theory of knowledge, what other theories of knowledge exist, and what would be their implications for understanding "relevance"? And he did not consider if the other views he presented could possibly be better interpreted as alternatives to the pragmatic view of knowledge. I believe an answer to this question uncovers theories of relevance at a deeper level and does so in a more satisfactory way than the system–user dualism of the received view.

Consider once more the quote from Borlund: "Basically, the concept of relevance can be divided into two main classes of relevance [. . .] (1) objective *or* system-based relevance; and (2) subjective *or* human (user)-based relevance" (Borlund, 2003, p. 914). Problems of objectivity and subjectivity are closely related to theories of knowledge. If we ignore the system–user dualism, could the basic classification of theories of relevance be (1) objective theories of relevance and (2) subjective theories of relevance?

I believe that such an outlook will turn out to be fruitful, but also difficult and full of paradoxes. The idea that each individual user has private relevance criteria can be seen as a kind of subjectivism,[48] while the idea that experts have qualified opinions about relevance (in relation to given tasks) can be seen as a kind of objectivism. This is complicated by a range of issues, however:

(1) Also the opinions of experts must be considered subjective.
(2) Subjective and objective are not necessarily opposites.[49]
(3) Given that expert opinions also are "subjective," what is the meaning of "objective relevance"?[50]
(4) The user-oriented tradition in relevance research tries to discover common patterns among users (i.e., "objective patterns").

DOI: 10.1002/asi

TABLE 3. The received view and a reinterpreted view.

"Received view" of categories of relevance	Suggested categories of relevance theories
Systems or algorithmic relevance "[S]ystem relevance, that is, system performance" (Saracevic 2008, p. 768). "The Cranfield model does not deal with dynamic information needs but treats information needs as a static concept entirely reflected by the user request and search statement. Further, this model uses only binary, topical relevance ignoring the fact that relevance is a multidimensional and dynamic concept." (Borlund, 2000a, abstract).	Universalism/objectivism A document or a piece of information is either relevant or is not relevant (or has a given degree of relevance), although assessors vary widely in their judgments. Expert assessment is generally needed to determine relevance, but is seen as neutral and objective (although with great variations). One system or algorithm may be optimal for all purposes and queries. Relevance is thus not seen as dependent on the user's theories, interests or different tasks and goals. Positivist assumptions
User oriented and cognitive relevance "The subjective class of relevance may, as a generic concept, refer to the aboutness, usefulness, usability, or utility of information objects in relation to the fulfillment of goals, interests, work tasks, or problematic situations intrinsic to the user. It is context dependent." (Borlund, 2003, p. 915). "The user employs the intellectual interpretation of a given work task situation as a platform (CW) for the relevance judgement." (Borlund, 2003, p. 916). "These variables can be standardized by having the same set of *simulated work task situations* searched by several test persons." (Borlund, 2003, p. 923).	Individual subjectivism Relevance is an individual/subjective experience and therefore "real user" assessment is needed. Paradoxically universal patterns underlying individual behavior are assumed to exist and form the basis for designing information systems. One system or algorithm is assumed to be optimal for all purposes and queries. There is a contradiction in the assumption that the user is seen as a subjective interpreter, but the researcher studying the users is assumed to be an objective investigator (positivist assumptions underlying user studies)
Subject knowledge view of relevance (Domain-oriented view of relevance) This view is not a part of the received view, but was originally suggested by Foskett (1972) and Saracevic (1975), was ousted and re-established by, among other works, Hjørland & Albrechtsen (1995). This view is not about "the system" (understood, for example, as library systems or computer systems), but is about the accumulated human knowledge or "the information ecology." The subject literatures form an important focus of this view (including bibliometric studies).	Pragmatism A document or a piece of information is either relevant or is not relevant to a given task relative to goals, values and interests. A document may be relevant even if nobody (yet) thinks so (objective relevance). It may thus turn out to be relevant in the future. Expert assessment is needed, but experts may have different interests, goals and values (also different from those of users), and they are not seen as neutral or objective assessors. Also, the opinions of experts change when they change theories. A given system may be optimal for some queries but not for all. Different interests and perspectives need different kinds of algorithms and systems.

(5) The user-oriented tradition in relevance research is itself based on an idea of "objective" study of users' behavior (related to behaviorism and positivism).

As I see it, both "the Cranfield tradition" and "the cognitive user-oriented tradition" are based on norms corresponding to the ideals of logical positivism in that they believe there are objective, neutral ways to construe optimal information systems and that "one size fits all." They are thus both opposed to the view of pragmatism that different kinds of queries, interests, users, and paradigms need different information systems: That no "one size" fits all. The Cranfield experiments may have demonstrated correctly that traditional systems such as UDC on average perform worse than free text searching in titles and abstracts, but I believe that traditional systems based on "controlled vocabularies" may be better for certain kinds of queries. Neither "the system tradition" nor "the user-oriented tradition" has helped us to identify the kind of solutions that are best suited to specific kinds of queries. They have both failed in just trying, on a statistical basis, to discover an average-based solution. Because no information need is ever "average," all information needs suffer when they are averaged. From the pragmatic perspective the important thing is not whether statements of relevance are subjective or not (they always are[51]). The important thing is to consider what kind of subjectivity should be preferred and which should be avoided?

Table 3 presents in the left column the units of "system," "user," and "domain" as the seats for studying relevance in the three corresponding views. In the right column an interpretation based on corresponding epistemological views is suggested.

This paper has thus defended a view of relevance as based on a pragmatic theory of knowledge and has suggested that the received view could be interpreted as variations of positivist epistemology, one focusing on expert opinion, the other on users' opinions. In 1975 Saracevic found "the subject knowledge view" to be the most important perspective of

DOI: 10.1002/asi

relevance, fundamental to all other views of relevance. This paper has confirmed this assessment. The importance and urgency of the work on that view is today more needed than ever.

Acknowledgments

I thank Associate Professor Jan Nolin for many valuable comments and suggestions, including an improvement of the discussion of Al Gore (2006) and of reflexivity. Jan also suggested the present title of the paper. Thanks to Dyveke Sijm for valuable suggestions, in particular for criticizing an imbalance in the first version of the paper. Thanks to Jeppe Nicolaisen for, among other things, proofreading. Thanks to two anonymous reviewers for constructive and detailed criticism.

Endnotes

[1]Saracevic (1975) is today (July 2009) the most cited paper in Social SciSearch about relevance (s relevance/ti and Sc=Library. Cited 59 times corresponding to 18.6% of all citations); Schamber, Eisenberg, and Nilan (1990) is the next most cited paper, with 53 citations corresponding to 16.7%.

[2]Tom Wilson (1981) and Sundin & Johannisson (2005) also describe the need for the social dimension in the understanding of relevance and information needs within LIS. Wilson (1981, p. 9) stated: "Because the situations in which information is sought and used are social situations, however, purely cognitive conceptions of information need are probably inadequate for some research purposes in information science, but not for all." Sundin & Johannisson wrote: "Despite Wilson's [1981] argument that was presented so many years ago, social aspects of information needs and relevance assessments have not been explored to any great extent. As a symptomatic indication of this state of affairs, individual aspects are very prominent when Donald Case [2002] in his recent book summarizes IS literature. For example, Case's book does not include any discussion at all of the social aspects of relevance assessments. This exclusion is not stated explicitly. Still, different epistemological approaches always—explicitly or implicitly—mediate specific views on how information needs are formed and satisfied by information, which is assessed as relevant from this specific viewpoint" (Sundin & Johannisson, 2005, p. 109).

[3]'System's view' (with apostrophe, possessive) as well as "system view," among other expressions have been used in the literature. Saracevic (1975, p. 326) used the term "the System's View of Relevance," while Borlund (2003, p. 914) used "system-based relevance" and "System or algorithmic relevance," and Vakkari & Hakala (2000, p. 540) used the term "[t]he system's relevance."

[4]It should be mentioned, however, that a related dichotomy was also expressed by Cranfield researchers: "These two types of relevance are called 'user relevance' and 'stated relevance.' The former can only be decided by the questioner himself, but 'stated relevance' can be determined [. . .] by anybody with reasonable knowledge of the subject field" (Cleverdon & Keen, 1966, pp. 256–257).

[5]The famous measures "recall" and "precision" are probably the most important characteristics of the Cranfield tradition, and these measures have been criticized for being static and objective. They are, however, dependent on how relevance is interpreted. They are as dynamic and as subjective as is the underlying interpretation of relevance!

[6]A colleague suggested to me that the early primitive computers sometimes made simple mechanical errors and that this was the origin of this understanding of relevance. If one searched for, for example, "user," the computer did not always return documents even if they were indexed by the term "user." I have since experienced the same kind of error on PsycINFO in Dialog. Evidently, during a reloading, identity numbers were mixed up so that terms in the inverted files did not match the correct records any longer. Although this may or may not be the etymological explanation of "systems relevance," it is totally irrelevant for present-day discourses of the meaning of "relevance."

[7]Some algorithms are used in systems based on "relevance feedback" or "find more like this." Such solutions are seen by some researchers as promising (e.g., Sparck Jones, 2005). In Hjørland (2008a) I argued against this by showing that in order to provide adequate feedback, users must already possess knowledge closely related to the information for which they are searching. Mechanisms behind "more like this" are just algorithms with the same kind of problems as already discussed. In addition it is a principal weakness that they provide *unspecified* "related documents" in contrast to systems where the user can choose, e.g., documents by the same author, containing the same words in the titles (or elsewhere), documents co-cited with another given document, etc. (i.e., well-defined and specified relations).

[8]Databases are different. When a given element (e.g., abstracts) is missing, it cannot be used. This is a definite limitation. If a user does not know how best to use the system, a subjective limitation is at play.

[9]In this article "information science" and "library and information science" are considered synonyms.

[10]In some databases, for example, OPACs, some of the data may be provided by general librarians, others by subject specialists. In reality even a given record in a database can thus be considered part of different information systems. As mentioned in the text, KOS also may be regarded as relatively independent kinds of "systems."

[11]An important aspect of databases is that terms are merged and taken out of context, and therefore parts of their meaning are lost, but may be reestablished. This is a core—but rather neglected—problem in retrieval interactions.

[12]Highest level (1) of relevance contains a complete answer to the question; level (2) references are relevant in the sense that "the lack of which either would have made the research impracticable or would have resulted in a considerable amount of extra work." Level (3) is "References which were useful, either as a general background to the work or as suggesting methods of tackling certain aspects of the work," level (4) is "References of minimum interest, for example, those that have been included from a historical viewpoint," and level (5) is "References of no interest" (Cleverdon, 1970, p. 3).

[13]Borlund (2000, p. 49–50) pointed out that relevance assessments "in the cognitive sense" are subjective, while also the Cranfield experiments "cannot be strictly objective." I shall return to the question of whether subjectivity in assessment is primarily associated with one of the sides of the system-user dualism.

[14]Ellis (1996, p. 28) wrote that "Cleverdon's choice of user relevance rather than stated relevance represented an attempt to simulate in the experimental procedure a position as near as possible to the real life situation of a user coming to an information retrieval system with a real information need."

[15]"With lower expertise this overlap [in relevance agreements] drops dramatically." Saracevic (2008, p. 776) further wrote that high expertise provides a high level of overlapping relevance assessments while low subject expertise provides a low overlap in assessments. This difference is declared "dramatic." Relevance judgments are thus less messy when done by experts.

[16]Sometimes the term "cognitive" is also added to this view of relevance (see Borlund, 2000, p. 58). I have not in this paper made any attempt to distinguish those studies that are "cognitive" from those that may be based on other approaches to studying users.

[17]Schamber et al. (1990, p. 768) also discuss this dichotomy and also quote Winograd & Flores (1986). Mizzaro (1998, p. 317) found that: "It is short-sighted to speak merely of 'system relevance' (the relevance as seen by an IR system) as opposed to 'user relevance' (the relevance in which the user is interested)" and that "'topicality' (a relevance for what concerns the topic component) is conceptually different from 'system relevance'."

[18]A system is meant to be designed for users, and users are in our context only relevant in relation to information systems and services. To believe it is possible to study them separately and provide useful knowledge is therefore mistaken.

[19]Hjørland (1997) established the following connection between the concepts "relevance" and "information need": To say that a document is relevant (to somebody, in a given situation) is the same as saying that the person

DOI: 10.1002/asi

needs that document. This is theoretically associated with a critique of Taylor's (1968) theory of the development of information needs as an internal psychological process.

[20]This paper has only been cited nine times in Web of Science (as of July 11, 2009). It seems much overlooked. The same is the case with Fugmann (1973), which had only been cited 12 times.

[21]This section was first written by the present author in a Web-page as a contribution to a collective book about Information Science to be published by ASIST. Retrieved September 28, 2009, from http://www.asis.org/wiki/pubs/index.php?title=2.B._Information_Needs&direction=next&oldid=2633

[22]Perhaps, as in Borlund, 2000, it is ensured that the students find the task "realistic." In this case there is also a dichotomy between general information sources (which it is believed that anybody may represent) and specialized sources which have to involve specialists (p. 134). This is, however, not examined or discussed in depth.

[23]Although it should be recognized that Schamber et al. (1990) emphasize the situational nature of information needs and do ask for investigations in different situations, their research questions seem in a way to contradict this. For example (p. 773), "can users recognize, or do they recognize, these clues?" It should be rather obvious that important principles are taught in higher education in courses in methodology. Students who have taken such courses should be able to recognize the clues, while students without such knowledge should not. Another example: Can users rely on the titles of documents to determine their relevance (i.e., titles as clues for relevance)? A very important part of the answer is that some titles are more reliable than others and that there are differences between disciplines and other "cultures," and therefore research is needed about the norms of the design of titles (and other document parts) in different domains. If we follow that idea, we have to leave "the user-based" view of relevance and study document architectures (seen by some as "the systems view").

[24]Tang & Solomon (1998) is a concrete study because it is about one person's specific search. However, the person is an anonymous student and we have no information that makes it possible for us to examine how she evaluated the references, whether this evaluation was qualified or not—or if her thesis was a good one or not. Perhaps she would change relevance criteria if she were to remain in this research area for much longer? Retrieval systems should be based on what, in the longer term, supports good theses and good research.

[25]To consider something an "error" may be regarded as "positivist" because it is assumed that there is one truth and that the researcher can identify that truth. Nonetheless, errors exist, but in many instances the issue is to consider the relative benefits and drawbacks of different solutions, why it is better to say that concrete failure analysis provides an interpretation of whether some specific terms should be regarded as fruitful and others should be regarded as unfruitful in indexing and searching.

[26]If the reason is that information scientists lack concrete knowledge, and if the argument made here is correct, then the implication is that the education of information scientists must be changed in order to provide necessary subject knowledge. What I am opposing is thus the tendency to provide courses in empirical research methods and to give the impression that these are sufficient for studying users and that taking issue with the concrete issues is outside the scope of information science. Because this has been a dominating tendency, studies of "human information behavior" may be in a state of crisis (cf. Cronin, 2009, who wrote: "A great deal has been written on the subject of information seeking over the years [. . .] but there is a regrettable lack of cumulation and coherence.")

[27]This may at first seem to be in contradiction with the criticism of relevance research being based on universalistic views. However, this is not the case. Different domains apply different criteria of relevance. When this is not specified, claimed findings and principles should apply to both researchers and the "users" being studied.

[28]The relevance of full text documents should be distinguished from estimating the relevance of a document given limited information about that document. The first is about the relevance of a document to a given task, the second is about the quality and utility of a bibliographical record (a "surrogate") in the absence of, or as supplement to, the full text. This should not be mixed together and it should be kept in mind that a core objective in library and information science is to optimize bibliographical records, and therefore a main goal is to determine their quality and what kind of value-added information bibliographical records should provide. There are different norms for composing documents that may pose different demands on records from different domains. Bibliographical records have different quality and different utility according to what else is available about the documents they index. We need to have some knowledge of this, not just some user's evaluation based on records the quality and function of which are not examined. This turns the study away from users and toward document composition, bibliographical records, and subject access points (i.e., "the information ecology"), cf. Hjørland & Kyllesbech Nielsen (2001).

[29]I am not implying that users should have a strong authoritative attitude, far from it. Patient groups and information services such as Wikipedia are certainly able to improve patients' and their relatives' knowledge. The philosopher John Dewey argued that general education was the best weapon against making people too dependent on experts. One of the anonymous referees kindly provided the following information: "A good read on this point—and the bonus is, it's hilarious—is the book Bad Science by Ben Goldacre [2008]. He has a website [http://www.badscience.net], but on the specific point about expert medical knowledge vs. fraudulent claims to expert medical knowledge, and claims to expert medical knowledge based on acquaintance with media reports of medical stories, the book is very useful. It would deepen the commentary that follows this, e.g., on quackery."

[30]Although it might be uncomfortable to problematize users' view of their information needs. It may feel like "besserwissen" (know-it-all).

[31]There have been some studies on the scientific consensus about climate change and Al Gore based his film on Oreskes (2004). A more recent study is Doran & Kendall Zimmerman (2009), which seems to provide a more nuanced view: "76 out of 79 climatologists who 'listed climate science as their area of expertise and who also have published more than 50% of their recent peer-reviewed papers on the subject of climate change' believe that mean global temperatures have risen compared to pre-1800s levels, and 75 out of 77 believe that human activity is a significant factor in changing mean global temperatures. Among all respondents, 90% agreed that temperatures have risen compared to pre-1800 levels, and 82% agreed that humans significantly influence the global temperature. Petroleum geologists and meteorologists were among the biggest doubters, with only 47 percent and 64 percent, respectively, believing in human involvement." This study is thus more in line with the domain-analytic hypothesis that views change according to domains but it does not change the conclusion about reservations in user-based relevance assessments.

[32]The user may of course not be able to understand scientific journals and prefer other sources. This is the kind of "subjective relevance" which is often studied. The underlying idea is problematic, however. If the most valid sources are considered nonrelevant the user may be subject to quackery. The point is that it requires subject knowledge to evaluate the relevance of information sources.

[33]The word "topicality" is often used in the meaning of "currency," i.e., "the topicality of the problem of the person in Islamic thought" or "the library had to discard books that had lost their topicality." In relevance discourses in information science, the meaning of "topicality" is mostly "being on the topic," i.e., about subject relations: "Topicality measures the 'aboutness' of a document to the topic area suggested by a query" (Xu & Yin, 2008, p. 202) or "Topicality, the relation of a document to the topic of a user's query" (Janes, 1994, p. 161).

[34]This aspect is recognized by Borlund: "In Saracevic's subjective topical-like relevance category the concept of topic is understood as aboutness, not contents, i.e., an intellectual assessment of how an information object corresponds to the topical area required and described by the request(s) for information. This relevance type is consequently not entirely based on the relationship between a query representation and a retrieved information object. An observer, either an assessor or a user, makes the judgment. It is this kind of subjective relevance assessment we ascribe to individual assessors who participate in common IR experiments like TREC, [note omitted] although their judgments traditionally are intended to be of an objective nature. To distinguish the subjective type of topical oriented relevance form [sic] algorithmic relevance we name this type of relevance:

DOI: 10.1002/asi

intellectual topicality (see Figures 1 or 2). Confusion exits in the literature when the topicality label of relevance is used to refer to either algorithmic relevance or intellectual topicality, and sometimes to even both of them at the same time, as in the case of Schamber et al. (1990)" (Borlund, 2003, p. 915).

[35]Beginning in 1964 in Pittsburgh, information science gradually became institutionalized in schools of librarianship. This is the origin of the term "library and information science," LIS. Information science originated in science, but schools of Library and Information Science have typically had difficulties recruiting students with a background in science. This development may have changed the field, not by arguments, but simply by changing interests and qualifications. The emergence of the user-based view of relevance may be related to such institutional factors.

[36]An important link between "the subject literature view" and "the subject knowledge view" is the concept of "literary warrant" (due to Hulme, 1911). In a broad interpretation of Hulme, we may say that criteria of relevance should be based on the literature rather than suggested without consulting it. Although some may mistakenly confuse this with "the system's view," it is clearly a third position in relation to the systems view and the user-based view.

[37]By defining relevance in relation to tasks, I am in agreement with a version of what has been termed "situational relevance," which was first formulated by Patrick Wilson. While there seem to be conflicting interpretations of this concept in the literature, I agree with the following quote: "As with logical and evidential relevance, the recognition that an item is relevant is not an automatic or mechanical procedure, and may be even beyond my abilities. That an item is situationally relevant is a logical, not a psychological, fact" (Wilson, 1973, p. 464).

[38]What makes this much more complex is that theories are seldom totally given up (falsified). A theory which for a period seemed abandoned often turns up later based on new arguments and findings.

[39]White (1992, p. 146) introduces the terms "ministerial relevance" and "magisterial relevance." "Ministerial relevance" means that somebody finds something relevant, but if another person does not share his belief, he will not consider using it. "Magisterial relevance" means that somebody has the power to determine what others should use. A teacher may, for example, decide what should be read. That means that a student's "own" information need is not considered. These two concepts clearly demonstrate the social nature of relevance.

[40]Saracevic (1975, p. 335) wrote: "Pragmatism has been a subject for serious criticism in philosophy." Yes it has, but so have the alternatives. We have to develop an understanding that is best suited for the tasks it has to do for us. If somebody can provide an argument for a position that is better than pragmatism, then I have to reconsider my position. So far, however, I consider pragmatism the best alternative. It should be considered, however, that there are different versions of pragmatism and that pragmatism may still be seen as an emerging theory.

[41]Saracevic (2007, p. 1918) uses the term "messy" about the notion of relevance. The messiness of relevance is at play at two levels. It is at play in all concrete determinations of relevance. Here they depend on theoretical clarity and subject qualifications. If a field is theoretically muddled, then also relevance assessment is muddled; if the assessor is unqualified, the relevance is wrong or vague. At the theoretical level this messiness can be considerably reduced if the perspective is changed from the individual perspective of the cognitive view to the social perspective of domain analysis.

[42]Relevance may be studied, for example, by considering criteria of good book reviews or by considering the criteria used by journals such as JASIST and Journal of Information Science in the instruction of referees. The latter journal asks the referees to consider, among other things:

- Does the paper argue an interesting case well?
- Is it a new and original contribution?
- Is any significant argumentation or evidence missing from the paper?
- Are the methods sound and adequately described?
- Are the conclusions and interpretations sound and justified by the data?
- Is the title accurate and succinct?

[43]If a document challenges the view of the user (or is based on problematic scholarly arguments and research methods), s/he may consider its argument and decide to (1) change view, (2) argue against it, or (3) just ignore it. If s/he decides to argue with it, it becomes relevant because this argument is now a new task in which the professional engages himself or herself. If s/he decides not to argue with it in the present paper, but perhaps to do so later, then it is not considered relevant for the present task, but for a potential future task.

[44]Perhaps there are so many ways and such a lack of consensus that there is very little help to get: This would be a problematic situation for our civilization, which is very much dependent on knowledge production and higher education, which is again dependent on the idea that it is possible to develop knowledge in ways that can be learned, i.e., is based on a distinction between expert and novice.

[45]Although views related to logical positivism still exist in some approaches, e.g., in "evidence-based medicine."

[46]If there are conflicts between "classical source criticism" and "feminist epistemology," a negotiated position may be developed or the historian will choose based on his/her present understanding.

[47]Although some positions have much more consensus than other positions, due to more well-defined principles and thus relevance criteria.

[48]If not solipsism. See Thornton (2006).

[49]"We shall not dwell at such length on the notion of subjectivity, insofar as it refers to the opinions, beliefs, and feelings of conviction of this or that individual. Let us mainly note that this is not in any way the logical opposite of objectivity. People said to be 'reasonable' or 'sensible' will often give their (subjective) agreement to a well-corroborated (objective) statement such as 'when an apple becomes detached from a tree, it falls down and does not fly towards the stars.' In that sense, obviously, any probabilistic statement, insofar as some individual expresses his support for it, can always be said to be subjective. But this does not exclude a priori its objectivity. An objective law, such as the law of universal attraction, insofar as I believe it to be 'true' can also be said to be subjective, since it does, in fact, represent my personal opinion" (Matheron, 1989, pp. 26–27).

[50]Objective relevance is what may serve as a tool even if nobody knows that or believes so. Uranium was a potential energy source before radioactivity was discovered. Many documents in libraries that nobody today considers relevant may be assumed to be relevant in the future. As human knowledge develops, so does our knowledge of what is relevant, whether we are speaking of material things, ideas, or documents. This is very much a reality in big libraries and databases. If documents had been culled because their relevance was not foreseen, many scholars would have suffered because views of what is relevant changes, "Today's dust is tomorrow's gold."

[51]It may seem strange (and offensive) to say that competent people's knowledge is "subjective." It should not be understood as "besserwissen," but as a principle based on fallibilism (known from, among others, Peirce and Popper) indicating that knowledge claims are invariably vulnerable and may turn out to be false (not to be confused with skepticism). In this perspective all knowledge should be understood as subjective knowledge claims, and all statements about relevance should also be understood as subjective. However, some knowledge claims and some relevance judgments are very well based, so well that fallibilism seems very hypothetical.

References

Abrahamsen, K.T. (2003). Indexing of musical genres. An epistemological perspective. Knowledge Organization, 30(3/4), 144–169.

Applegate, R. (1993). Models of user satisfaction: Understanding false positives. RQ, 32(4), 525–539.

Bloor, D. (1976). Knowledge and social imagery. London: Routledge & Kegan Paul.

Borlund, P. (2000). Evaluation of interactive information retrieval systems. (Doctoral dissertation, Åbo Akademi University).

Borlund, P. (2003). The concept of relevance in IR. Journal of the American Society for Information Science and Technology, 54(10), 913–925.

Case, D. (2002). Looking for information: A survey of research on information seeking, needs, and behavior. Amsterdam: Academic Press.

Cleverdon, C.W. (1970). The effect of variations in relevance assessments in comparative experimental tests of index languages. (Cranfield Library Report No. 3). Cranfield, UK: Cranfield Institute of Technology.

DOI: 10.1002/asi

Cleverdon, C.W. (1971). Design and evaluation of information systems. Annual Review of Information Science and Technology, 6, 41–73.

Cleverdon, C.W., & Keen, M. (1966). Aslib Cranfield research project—Factors determining the performance of indexing systems. Volume 2, Test results. Cranfield, UK: ASLIB Cranfield Research Project. Retrieved July 22, 2009, from https://dspace.lib.cranfield.ac.uk/bitstream/1826/863/2/1966e.pdf

Cleverdon, C.W., Mills, J., Keen, M. (1966). Factors determining the performance of indexing systems. Volume 1, Design: Part 1. Cranfield, UK: ASLIB Cranfield Research Project. Retrieved July 22, 2009, from http://hdl.handle.net/1826/861; Part 2, Appendices. Retrieved July 22, 2009, from https://dspace.lib.cranfield.ac.uk/handle/1826/862

Coiera, E.W., & Vickland, V. (2008). Is relevance relevant? User relevance ratings may not predict the impact of Internet search on decision outcomes. Journal of the American Medical Informatics Association, 15(4), 542–545.

Cooper, William S. (1971). A definition of relevance for information retrieval. Information Storage and Retrieval, 7, 19–37.

Cronin, B. (2008). The sociological turn in information science. Journal of Information Science, 34(4), 465–475.

Cronin, B. (2009). Introduction. Annual Review of Information Science and Technology, 43, VII–X. Medford, NJ: Information Today.

Doran, P.T., & Kendall Zimmerman, M. (2009). Examining the scientific consensus on climate change. Eos, Transactions, American Geophysical Union, 90(3). Retrieved August 23, 2009, from http://tigger.uic.edu/≈pdoran/012009%5FDoran%5Ffinal.pdf

Ellis, D. (1996). The dilemma of measurement in information retrieval research. Journal of the American Society for Information Science, 47(1), 23–36.

Foskett, D.J. (1972). A note on the concept of 'relevance.' Information Storage and Retrieval, 8(2), 77–78.

Frohmann, B. (1990). Rules of indexing: A critique of mentalism in information retrieval theory. Journal of Documentation, 46(2), 81–101.

Fugmann, R. (1973). On the role of subjectivity in establishing, using, operating and evaluating information retrieval systems. Treatise II on retrieval system theory. Information Storage and Retrieval, 9(7), 353–372.

Gerhart, S. (2004). Do web search engines suppress controversy? First Monday, 9(1). Retrieved September 28, 2009, from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1111/1031

Goldacre, B. (2008). Bad science. London: Fourth Estate.

Guggenheim, D. (Director). (2006). An Inconvenient Truth [Motion picture]. United States: Paramount Classics.

Hildreth, C.R. (2001). Accounting for users' inflated assessments of online catalogue search performance and usefulness: An experimental study. Information Research, 6(2). Retrieved October 20, 2009, from http://InformationR.net/ir/6-2/paper101.html

Hjørland, B. (1997). Information seeking and subject representation. An activity-theoretical approach to Information Science. Westport, CT, and London: Greenwood Press.

Hjørland, B. (2000). Relevance research: The missing perspectives: "Nonrelevance" and "epistemological relevance." Journal of the American Society for Information Science, 51(2), 209–211.

Hjørland, B. (2002). Epistemology and the socio-cognitive perspective in Information Science. Journal of the American Society for Information Science and Technology, 53(4), 257–270.

Hjørland, B. (2008a). Core classification theory: A reply to Szostak. Journal of Documentation, 64(3), 333–342.

Hjørland, B. (2008b). Source criticism. Wikipedia. The free encyclopedia. Retrieved July 15, 2009, from http://en.wikipedia.org/wiki/Source%5Fcriticism

Hjørland, B. (2009). Concept theory. Journal of the American Society for Information Science and Technology, 60(8), 1519–1536.

Hjørland, B., & Albrechtsen, H. (1995). Toward a new horizon in Information Science: Domain analysis. Journal of the American Society for Information Science, 46(6), 400–425.

Hjørland, B., & Kyllesbech Nielsen, L. (2001). Subject access points in electronic retrieval. Annual Review of Information Science and Technology, 35, 3–51.

Hjørland, B., & Nicolaisen, J. (2005). The epistemological lifeboat. Epistemology and philosophy of science for Information Scientists. Retrieved July 15, 2009, from http://www.db.dk/jni/lifeboat/default.asp

Hjørland, B., & Sejer Christensen, F. (2002). Work tasks and socio-cognitive relevance. A specific example. Journal of the American Society for Information Science and Technology, 53(11), 960–965.

Holland, K. (2005). Relevance (debate). In B. Hjørland & J. Nicolaisen (Eds.), The epistemological lifeboat. Retrieved July 6, 2009, from http://www.db.dk/jni/lifeboat/info.asp?subjectid=176

Hulme, E.W. (1911). Principles of book classification. Library Association Record, 13, 354–358, 389–394, 444–449.

Introna, L.D., & Nissenbaum, H. (2000). Shaping the Web: Why the politics of search engines matter. The Information Society, 16(3), 169–186. Retrieved June 7, 2007, from http://www.nyu.edu/projects/nissenbaum/papers/searchengines.pdf

Janes, J.W. (1994). Other people's judgments. A comparison of users' and others' judgments of document relevance, topicality, and utility. Journal of the American Society for Information Science, 45(3), 160–171.

Kann-Christensen, N., & Andersen, J. (2009). Developing the library: Between efficiency, accountability and forms of recognition. Journal of Documentation, 65(2), 208–222.

Kuhn, T.S. (1962, 1970, 1996). The structure of scientific revolutions. Chicago: University of Chicago Press.

Matheron, G. (1989). Estimating and choosing. An essay on probability in practice (A.M. Hasofer, Trans.). Berlin, Germany: Springer.

Mizzaro, S. (1997). Relevance: The whole history. Journal of the American Society for Information Science, 48(9), 810–832.

Mizzaro, S. (1998). How many relevances in information retrieval? Interacting with Computers, 10(3), 303–320.

Nicolaisen, J. (2009). Compromised need and the label effect: An examination of claims and evidence. Journal of the American Society for Information Science and Technology, 60(10), 2004–2009.

Oreskes, N. (2004). Beyond the ivory tower: The scientific consensus on climate change. Science, 306(3 Dec.), 1686. Retrieved August 23, 2009, from http://www.sciencemag.org/cgi/content/full/306/5702/1686

Ørom, A. (2003). Knowledge organization in the domain of Art Studies—History, transition and conceptual changes. Knowledge Organization, 30(3/4), 128–143.

Rees, A.M., & Schultz, D.G. (1967). A field experiment approach to the study of relevance assessments in relation to document searching, 2 vols. Cleveland, OH: Center for Documentation and Communication Research, School of Library Science, Case Western Reserve University.

Ryle, G. (1949). The concept of mind. London: Hutchinson & Co.

Saracevic, T. (1970). The concept of "relevance" in information science: A historical review. In T. Saracevic (Ed.), Introduction to information science (pp. 111–151). New York: R. R. Bowker.

Saracevic, T. (1975). Relevance: A review of and a framework for the thinking on the notion in information science [in 2007 termed "Part I (of III)"]. Journal of the American Society for Information Science, 26 (6), 321–343. Retrieved September 28, 2009, from http://www.scils.rutgers.edu/≈tefko/Saracevic_relevance_75.pdf

Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. Journal of the American Society for Information Science and Technology, 58(3), 1915–1933. Retrieved June 12, 2009, from http://www.scils.rutgers.edu/≈tefko/Saracevic%20relevance%20pt%20II%20JASIST%20%2707.pdf

Saracevic, T. (2008). Effects of inconsistent relevance judgments on information retrieval test results: A historical perspective. Library Trends, 56(4), 763–783. Retrieved June 12, 2009, from http://www.scils.rutgers.edu/≈tefko/LibraryTrends2008.pdf

Schamber, L. (1994). Relevance and information behavior. Annual Review of Information Science and Technology, 29, 3–48.

Schamber, L., Eisenberg, M.B., & Nilan, M.S. (1990). A re-examination of relevance: Toward a dynamic, situational definition. Information Processing & Management, 26(6), 755–776.

Soergel, D. (1976). Is user satisfaction a hobgoblin? Journal of the American Society for Information Science, 27(4), 256–259.

Soergel, D. (1985). Organizing information: Principles of data base and retrieval systems. Orlando, FL: Academic Press.

Sparck Jones, K. (2005). Revisiting classification for retrieval. Journal of Documentation, 61(5), 598–601.

Sundin, O., & Johannisson, J. (2005). The instrumentality of information needs and relevance. In F. Crestani & I. Ruthven (Eds.), Information context: Nature, impact, and role: Fifth International Conference on Conceptions of Library and Information Sciences (CoLIS 2005). Lecture Notes in Computer Science, vol. 3507, S.107–118.

Tang, R., & Solomon, P. (1998). Towards an understanding of the dynamics of relevance judgments: An analysis of one person's search behavior. Information Processing & Management, 34(2/3), 237–256.

Taylor, R.S. (1986). Value added processes in information systems. Norwood, NJ: Ablex Publishing.

Thornton, S.P. (2006). Solipsism and the problem of other minds. The Internet Encyclopedia of Philosophy. Retrieved July 14, 2009, from http://www.iep.utm.edu/s/solipsis.htm

Vakkari, P., & Hakala, N. (2000). Changes in relevance criteria and problem stages in task performance. Journal of Documentation, 56(5), 540–562.

White, H.D. (1992). Literary forms of information work: Annotated bibliographies, bibliographic essays, and reviews of literatures. In H.D. White, M.J. Bates, & P. Wilson (Eds.), For information specialists: Interpretations of reference and bibliographical work (pp. 131–149). Norwood, NJ: Ablex Publishing.

White, H.D. (2001). Authors as citers over time. Journal of the American Society of Information Science and Technology, 52(2), 87–108.

Wilson, P. (1973). Situational relevance. Information Storage and Retrieval, 9, 457–471.

Wilson, T.D. (1981). On user studies and information needs. Journal of Documentation, 37(1), 3–15.

Winograd, T., & Flores, F. (1986). Understanding computers and cognition: A new foundation for design. New York: Addison-Wesley.

Xu, Y., & Yin, H. (2008). Novelty and topicality in interactive information retrieval. Journal of the American Society for Information Science and Technology, 59(2), 201–215.

DOI: 10.1002/asi