

Thus an enormous body of information science literature is based on work that *uses* relevance without thoroughly understanding what it *means*. A much smaller body of literature attempts, unsuccessfully, to reach a consensus on its meaning. In the late fifties, information scientists engaged in intense debates on theoretical and methodological aspects of relevance. In the sixties, two major experimental studies attempted to examine the factors that influence relevance judgments. In the seventies, a few ambitious attempts were made to establish a theoretical framework for relevance. Since then, overt interest in understanding the concept has seemed to subside. Yet serious questions about the nature of relevance remain. The need for a thorough definition is more pressing than ever, for at least three reasons:

1. Relevance is the measure of retrieval performance of all information systems, including full-text, multimedia, question-answering, database management, and knowledge-based systems. Increasingly complex systems are being developed that promise to serve users more effectively than ever. It is inevitable that these new systems, like systems since the library card catalog, will be evaluated (explicitly or implicitly) on the basis of human relevance judgments.
2. Among current developments are information retrieval systems that actually employ users' relevance judgments in the course of their operating processes. For example, relevance feedback mechanisms, in which users' relevance judgments are used to modify lists of documents as the search progresses, make users an integral part of the system. In such systems, relevance is no longer a *reactive* concept, to be used primarily in evaluation, but an *active* concept vital to the functioning of the system itself. However, without an understanding of what relevance means to users, it seems difficult to imagine how a system can retrieve relevant information for users.
3. Information scientists must finally establish a full theoretical and empirical understanding or definition of relevance as a fundamental concept, so that the discipline can move on to other matters.

Two key questions are at issue here: What is the meaning of relevance? and What role does relevance play in information behavior? This paper attempts to address these questions. It does so by reviewing the development of thought on the concept of relevance over more than 30 years; that is, since some of the first discussions on relevance appeared in print. The predominant model of information retrieval has been one of a communication process in which a message is transferred from a source (the system) to a destination (the user). But over the years, attention has gradually shifted from system mechanisms to the user; to a concept of relevance tied to human information behavior. This theme, of system and user, is one that cuts across all parts of the paper, although our emphasis throughout is more on *people* than on *things* such as information queries, documents, systems, or even information itself.

We begin, in part 2, by discussing issues concerning the definition of relevance, as well as existing system- and user-oriented definitions. Part 3 describes prominent multidimensional approaches to relevance, and part 4 cognitive and dynamic approaches. In part 5 we look at traditional assumptions underlying concepts of relevance and begin building a case for an alternative approach to a definition. Finally, in part 6, we suggest a dynamic, situational approach to defining relevance that focuses on users' perceptions. This approach differs philosophically from traditional approaches in that the user—regardless of system—plays the central role.

2. ISSUES CONCERNING DEFINITION

The matter of relevance has long sparked controversy in information science. In more than three decades of spirited debate and see-sawing opinion, the most consistent view has been that relevance pertains to more than system functions; that, ultimately, it pertains to people. Two arguments that surfaced during the first formal in-depth discussion of rele-

vance, at the International Conference for Scientific Information (ICSI) in 1958, have yet to be resolved. One argument focuses on what kinds of relevance are philosophically or empirically "best"; the other on what kinds of relevance can be measured.

Both arguments revolve around two broad notions of relevance, as presented by Vickery (1959a, 1959b) at the ICSI: *relevance to a subject* and *user relevance*. Relevance to a subject refers to matches between subject terms in queries and subject terms in documents, or what is commonly known as *topicality*. Topicality is often called a *system-oriented* view of relevance because it depends on factors related to systems.² User relevance, according to Vickery, is based ultimately on how far the user decides to pursue the search for information. This notion, obviously, depends on factors related to users and thus is called a *user-oriented* view of relevance.

"The consensus of the ICSI debate," Rees and Schultz (1967) wrote later, "was that (a) relevance is more than the operation of relating, performed internally within systems; (b) relevance is not exclusively a property of document content; (c) relevance is not a dichotomous decision; and (d) there is such a thing as 'user relevance' which *can* be judged." However, they said, "In general, relevance was viewed as the product of a match between the terms of a question and the terms of a document. Relevance was considered to be a property of the system, of its internal organization, and of the individual documents included. . . . In 1958, there was no explicit discussion of the subjectivity inherent in a human judgment of relevance. Also, the views held on relevance were based on intuitive inferences unsupported by generalizable experimental evidence" (pp. 7-8).

Cuadra and Katter (1967a) observed that the literature on system evaluation grew substantially in the decade following the mid-fifties, but "it does not appear that all this increased effort has yet resulted in clear-cut guidelines for system improvement, in tangible improvements in evaluation methodology, or in any widespread agreement among workers in this field as to fruitful areas for further inquiry." They also commented that "most of the workers engaged in system evaluation . . . have been aware of disagreement among [relevance] judges, but they have tended to consider such disagreement largely as an irritant, to be stamped out or bypassed as quickly as possible, rather than as a phenomenon worthy of interest in its own right" (pp. 12, 14).

All of these remarks might just as well have been made today. Most information scientists still seem to find it difficult to deal with theoretical issues related to relevance, while continuing to use relevance as a basis for system design and evaluation. In the remainder of this section we will briefly address the issue of measurement, then focus on the issue of different kinds of relevance.

2.1 *The measurability of relevance*

"The problems associated with measurement," said Ellis (1984), "have found expression most critically, and notoriously, on the question of relevance. In contemporary information retrieval research the problem of measurement is almost synonymous with the question of the validity of the employment of relevance as a performance criterion" (p. 26). The problem of validity has most often been expressed in terms of the objectivity (measurability) versus the subjectivity (unmeasurability) of different kinds of relevance, a matter that was debated hotly during the sixties and seventies (see Saracevic, 1970; Swanson, 1977). Bar-Hillel (1960), for example, criticized a topical, or system-oriented, definition of relevance as being logically unattainable, on the basis that distances between topics or between documents cannot be measured. On the other hand, Fairthorne (1963) criticized a user-oriented definition of relevance on the basis that it requires an impossible omni-

²Throughout this paper we have tried to use the same terminology as the writers whose work we are citing. However, terms such as *query*, *document*, and *system* have broader connotations now than they did when many of these papers were written. Therefore, let us say that *query* (or request) can refer to oral or written representations of information needs (or questions or problems); *document* can refer to representations of information in any form, including full-text document, document representation (e.g., citation, abstract), image, audio, multimedia, etc; and *system* can refer to bibliographic or document retrieval, full-text, image-based, multimedia, question-answering, database management, knowledge-based system, and so on.