DOCUMENTS, MEMORY INSTITUTIONS AND INFORMATION SCIENCE

BIRGER HJØRLAND
bh@db.dk

Royal School of Library and Information Science, Copenhagen

This paper investigates the problem of the labelling of the library, documentation and information field with particular emphasis on the terms 'information' and 'document'. What influences introduced the concept of 'information' into the library field in the middle of the twentieth century? What kind of theoretical orientations have dominated the field, and how are these orientations linked to epistemological assumptions? What is the implication of the recent influence of socially oriented epistemologies for such basic concepts in IS as 'information' and 'document'? The article explores these problems and advocates an approach with emphasis on documents and on the concept 'memory institutions' as generic terms for the central object of study.

1. INTRODUCTION

In 1997 the 'Royal School of Librarianship' in Copenhagen changed its name to the 'Royal School of Library and Information Science'. This is just one example of an influential trend towards an increasing use of the term 'information' in the library sector. Made topical by this change of name, this article aims to explore some of the influences and consequences of this trend and to provide some theoretical suggestions.

The term 'library science' goes back to the nineteenth century, where Martin Schrettinger introduced this concept in a textbook from 1808 [1, 2[1]], and in 1894 there existed a 'Department of Library Science' in Chicago. In the twentieth century this concept is used, among others, by Pierce Butler [3] and by S. R. Ranganathan [4, 5].[2] Although it is still used today (e.g. [6]) this term is by and large replaced by 'library and information science', 'LIS' (or often just 'information science', 'IS'). Thus, *Library Science Abstracts* changed its name to *Library and Information Science Abstracts* in 1969.

Another important term related to LIS is 'documentation'. Rayward writes [7, p. 238]: 'the term "documentation" is a neologism invented by Otlet to designate what today we tend to call Information Storage and Retrieval. In fact it is not too much to claim the *Traité* [8] as one of the first information science textbooks'.

---

[1]Schrettinger [1] is according to Kunze and Rückl [2, p. 267] the first person who used the word 'Bibliothekswissenschaft'.

[2]In other books, however, Ranganathan uses the term 'documentation'.

Meadows [9, p. 59] writes that the development of theory in information science originated in the needs of special libraries, which were recognised in the years between the two World Wars to differ from those of public, academic or national libraries.

> The main differences [between library science and documentation] were identified as lying in the areas of bibliography and what came to be called 'documentation'. Exactly what the differences between these new 'documentalists' and traditional librarians were was not altogether well defined. However, there was general agreement that documentalists were concerned not only with the physical handling of documents, but, to a much greater extent than traditional librarians, with the exploitation of the information contained in the documents. This practical thread generated some of its own theory, a noticeable example being Bradford's law of scattering.

Buckland [10, pp. 46–48] analyses important aspects of the history of documentation. Early in the twentieth century, the documentalists felt that there existed a need for a generic term, an expression for the objects covered by the activity of documentation. They included not only texts, but also natural objects, artefacts, models designed to represent ideas, and objects of art. The concept 'document' (or unit of documentation) was applied in a special meaning as a designation for informative physical objects. Buckland points out that the word 'document' originally meant a tool to teach or inform no matter if it were a lecture, an experience or a text. Limiting the meaning to objects carrying text came only at a later time.

Although the term 'documentation' is still in use[3], it has often been replaced by the word 'information'. One of the most pronounced changes towards 'information science' was the decision of the American Documentation Institute (ADI, founded in 1937) to change its name officially to the American Society for Information Science (ASIS) in 1968. As mentioned some institutions have preserved the concept 'documentation', and of course many institutions have preserved the concept 'library'. However, in recent years extremely few institutions have chosen the term 'documentation' as part of their name. One notable exception is in Tromsø, Norway, where 'Documentation Science' has been chosen as the name for a newly founded institute.

Two Scandinavian LIS schools, the Royal School of Library and Information Science in Copenhagen and the Institute of Documentation Science at the University of Tromsø, have thus chosen different names. What theoretical influences lie behind these choices? How do such different conceptions affect the content of the activities that are carried on? This article aims to present an interpretation of these views and to argue for an approach to IS with emphasis on such concepts as documents and 'memory institutions'[4] [11] as generic terms for the object of study in IS.

---

[3]The term documentation is, for example, used in *Journal of Documentation*.

[4]The term 'memory institutions' is taken from the Swedish information scientist R. Hjerppe [11].

2.   THE INFLUENCES OF INFORMATION TECHNOLOGY (IT)

In 1968 the American Documentation Institute changed its name to the American Society for Information Science (ASIS). A definition of IS was provided by ASIS in 1975:

> Information science is concerned with the generation, collection, organisation, interpretation, storage, retrieval, dissemination, transformation and use of information, with particular emphasis on the applications of modern technologies in these areas.
>
> As a discipline, it seeks to create and structure a body of scientific, technological, and systems knowledge related to the transfer of information. It has both pure science (theoretical) components, which inquire into the subject without regard to application, and applied science (practical) components, which develop services and products [12, p. 5; cf. 13, p. 3].[5]

The establishment of ASIS did not solve the theoretical problems concerning the subject matter of IS, which has always been an ill-defined concept. Schrader [14] studied about 700 definitions of 'information science' and its antecedents from 1900 to 1981 and found that:

> … the literature of information science is characterised by conceptual chaos. This conceptual chaos issues from a variety of problems in the definitional literature of information science: uncritical citing of previous definitions; conflating of study and practice; obsessive claims to scientific status; a narrow view of technology; disregard for literature without the science or technology label; inappropriate analogies; circular definition; and, the multiplicity of vague, contradictory, and sometimes bizarre notions of the nature of the term 'information' [14, p. 99].

The ASIS definition indicates that IS was born with the intention to contribute to the automation of processes in some specific areas such as libraries and bibliographical services. However, the sentence 'with particular emphasis on the applications of modern technologies in these areas' is a bit tricky. Computer science is the science about automation. ASIS defined IS as a specific content area 'with particular emphasis on the applications of modern technologies'. Libraries and many other institutions and services have traditionally been concerned with the collection, organisation, storage, retrieval and mediation of literature and documents. This is what 'library science' basically must be about. The application of the most adequate tools and modern technology to a given purpose should be taken for granted in all areas (including education, medicine, libraries and scientific communication). Just as we cannot think of two medical sciences, one technology oriented and one non-technology oriented, we cannot (or should not) think of two 'library sciences', one technology oriented and one non-technology oriented.[6] A science must be defined by its object, not by its tools.

---

[5]This definition quoted from [12] is rather close to a definition given by Borko [13]. This last reference was explicitly motivated by the change of name ADI made to ASIS in 1968.
[6]However, people in LIS can be more or less interested in different aspects; they can be more or less technology-oriented, practice-oriented, content-oriented, theory/philosophy-oriented, and so on.

The digitisation of information resources and services is an extremely important step in the development of (L)IS. The development of bibliographical databases (such as *MEDLINE* and *SciSearch*) and of database hosts such as Dialog (founded in 1971) can in my opinion not be overestimated. Important developments in online systems have been traced by Hahn [15]. They include the development of Boolean and proximity operators, masking of characters in suffixes or prefixes, numeric and date ranging, search term weighting, elimination of stop words, automatic incorporation of synonyms into search formulations, the capability of limiting a search to a specific field (author, geographic term, descriptor, title or other), citation searching, the ability to request the system to find 'more like this'; further improvements are: fuzzy search, ranking and relevance feedback on retrieved output, and highlighting (the capability of displaying the words in retrieved records that match the search terms), iterative search facilities (the capability to modify further the results from a previous search), canned or stored query (the capability to name a query and store it to be retrieved, executed and modified during a later session), vocabulary browse (that allows a user to display in alphabetical order the words from the document database), concept hierarchies and thesaurus expansion (that permit users to display hierarchical or conceptual relationships among terms based on predetermined subject relationships or on statistical relationships), etc. Such features in modern online systems represent very important and impressive progress.[7]

The opportunities for the manipulation of data that electronic retrieval systems provide are one thing; how such features can be utilised and their importance evaluated is quite another. When, for example, should a searcher use descriptors and when should he or she use citations? When should a search be restricted to titles or abstracts and when should full text be used? Is manual indexing really improving retrieval, and is it worth its costs? How can we define 'relevance', and how can we measure retrieval efficiency? We clearly lack theories to guide the investigations of such problems, which today seems to constitute a block to further advance. Very central for a necessary reorientation in IS are in my opinion both a new focus on meaning and a new focus on the social environments of users and systems. Van Rijsbergen [16, p. 194] has pointed out that the concept of meaning has been overlooked in information retrieval (IR), and this is the reason for the crisis in the field. He finds that the fundamental basis of all previous work – including his own – is wrong because it has been based on the assumption that a formal notion of meaning is not required to solve the IR problems. In other words: further progress seems to be much more dependent on issues related to 'cognitive' kinds of problems than to problems of a purely technological nature.

The idea that problems of a cognitive nature are essential for the further development of IS is certainly not new. This has been emphasised by researchers such as Taylor, Saracevic, Belkin, Ingwersen and many others. The new element in cognitive science and in IS is the view of cognitive processes. Where the traditional focus on cognition was influenced by rationalistic views of what goes on 'inside the head', more recent approaches to the study of cognition emphasise the role of cultural and domain specific factors in cognition [cf. 17–19]. Information

---

[7]All these achievements were done by people outside the library field.

technology is based on theoretical views such as cybernetics, information theory and systems theory. These theoretical views have also influenced our ways of thinking about people, knowledge, information and communication. The mathematical theory of communication [20] established a measurement of 'information' (the unit 'bit' was established later). It is something of an irony that IT has strongly influenced theoretical views in the 'cognitive sciences', and that today more and more researchers view this influence as a problematic theoretical framework.

IT has had a positive influence on IS because it changed the perspective of individual information services and made the foundation of a new perspective much more general, systems-oriented, flexible and research-based. When traditional catalogues in libraries were challenged by bibliographic databases with abstracts, the latter appeared to be far more satisfactory and efficient for the users. Further developments along this line (including full-text databases and the Internet) have questioned the future role of traditional libraries, archives and other kinds of 'memory institutions', and future-oriented researchers therefore often tried to contribute to solutions aimed at the digital future. In this process it is often forgotten that the core interest in IS is not IT in itself, but 'facilitating the effective communication of desired information between human generator and human user' [21, p. 22].

The conclusion of this section is that IT has influenced theory and terminology in the library/documentation/information field in different ways:

1. directly, by associating the scientific parts of library work with automation and 'information technology': the object of research in IS is library automation. Delimiting the field like this appears too narrow and such a focus neglects the investigating of mechanisms and dynamics underlying information use;

2. indirectly, by focusing on some specific theories or approaches, which showed themselves fruitful in computer technology but which turned out to be problematic as theoretical frames for LIS. An influential example is the attempt to apply Shannon's information theory to problems related to facilitating the effective communication of desired information between human generator and human user. Even among the researchers in IS who found Shannon's theory useless, certain ways of conceptualising problems remained influential. IT impacted on the theories in human psychology, where the so-called cognitive revolution established a new interdisciplinary research area called the cognitive sciences based on the assumption that the human brain is functionally identical to a computer. This trend also strongly influenced the view of users and of retrieval interaction in IS. In particular, information theory influenced the use of the information concept (at the expense of other concepts such as the concept of documents).

3. THE UNCLEAR MEANING OF THE TERM 'INFORMATION' IN IS

The meaning of a scientific concept is always determined by theoretical assumptions. Concepts have no meaning in themselves apart from theories or theoretical

assumptions. One can approach the term 'information' from the popular usage or the scientific usage of the term. It is, however, extremely important to realise the consequences of using different meanings of terms and to trace the often implicit theoretical influences and consequences behind given meanings. Analysing the implicit meanings of the term 'information' has important consequences for the kind of research and teaching which should be done in LIS. This paper will illustrate that.

In Section two above it was argued that the increasing use of the term 'information' in the library and documentation field is caused by direct and indirect influences from IT. Information is, however, an ambiguous concept. The ASIS definition of information science [12] stated that IS is concerned with the generation, collection, organisation, interpretation, storage, retrieval, dissemination, transformation and use of information, implying that information is something, which can be produced, stored, transformed and used. This definition has thus an implicit conception of information as either being physical units like documents or intangible units like decontextualised pieces of facts, opinion or ideas. Such an understanding is not in accordance with most recognised theories of information, which explicitly deny that information is a thing but see it as a change in the receiver's knowledge [e.g. 22, pp. 30–34]. What is processed in information systems[8] is only data (or potential information).

Because there is no need for both the term 'information' and the term 'document' unless they have different meanings, I shall concentrate my analysis and criticism on the widespread view that information is isolated pieces of facts, opinion or ideas that can be processed and managed in information systems. This view is, for example, used by *Encyclopædia Britannica* [23]: 'In popular usage, the term *information* refers to facts and opinions provided and received during the course of daily life'.

In the literature of IS there is a distinction between document retrieval and fact retrieval.[9] Distinguished researchers in IS have regarded the creation of fact retrieval systems as the ultimate goal of IS. Sparck Jones [24, p. 9], for instance, claims that 'we are concerned with access and, more materially, indirect access to the information the user wants: he wants the information in the documents, but the system only gives him the documents'. This statement represents a rather ordinary view with roots back to the foundation of documentation and information science. Some of Paul Otlet's basic ideas are described by Rayward [7] as 'the outmoded paradigm of nineteenth-century positivism':

---

[8]If the term 'information system' should be defended it should be seen as a system intended to *inform* actual or potential users about something.

[9]As response to a query a document retrieval system provides a list of references about the subject, which with a certain probability is supposed to contain the answer to the query or rather to reveal the present documented knowledge about the problem. Fact retrieval systems on the other hand are supposed to provide concrete answers to queries. If the query is: 'What is the definition of information science?', a document retrieval system such as *Library and Information Science Abstracts* (LISA) produces a long list of papers discussing this issue, whereas a fact retrieval system provides you with one selected definition, e.g. the ASIS definition cited earlier in this article.

> Otlet's concern was for the objective knowledge that was both contained in and hidden by documents. His view of knowledge was authoritarian, reductionist, positivist, simplistic – and optimistic! … It is merely a question of institutionalising certain processes for analysing and organising the content of documents. For him that aspect of the content of documents with which we must be concerned is facts. He speaks almost everywhere of *facts* … [7, p. 247].

Rayward can find the same view represented in modern IS:

> In describing the Xanadu Project, Nelson (1987) for example, in capital letters, says that it is 'just one thing: a new form of interconnection for computer files – CORRESPONDING TO THE TRUE INTERCONNECTION OF IDEAS which can be refined and elaborated into a shared network' (p. 143). These words and the sentiments that they both express and seem to imply could be, except for the term 'computer files,' Otlet's own. They suggest an atavistic positivist perspective that takes one by surprise [7, pp. 247–248].

In practice document retrieval systems coexist with systems that provide concrete answers. Directories, dictionaries, handbooks of chemical and physical constants, and many other kinds of reference works are examples of factographic works and databases that have important functions and exist side by side with bibliographic databases. However, I find it important to argue against the view that bibliographical databases or full-text databases should be less than ideal because '[the user] wants the information in the documents, but the system only gives him the documents' [24, p. 9].

The idea that bibliographical information systems should be reduced to fact retrieval systems is a problematic assumption. I agree with Rayward that this view is related to a kind of obsolete positivism. I also see this view as one reason for the use of the terms 'information' and 'information science'. Because it is of interest to know the source (in order to compare it with other sources and to evaluate its cognitive authority), document retrieval should *not* ideally be reduced to fact retrieval systems. The Danish linguist and information scientist Spang-Hanssen [25] once argued that what is fact is often debatable, and empirical facts may not have an all too safe future. In his opinion fact retrieval systems could be seen as systems which hide their information sources and keep them exclusively for themselves and only communicate some of their content.

The same kinds of reductionistic assumptions are also present in the information retrieval (IR) tradition starting with the Cranfield experiments in the 1950s and still continuing. This tradition has always concentrated on document/text retrieval, but very often researchers have hoped that it would be possible in one way or another to eliminate the concept of document/text and to store and retrieve just the facts or 'information' contained in the documents. Thus Ellis describes 'an anomaly' in IS: 'Brookes noted the anomaly could be resolved if information retrieval theory were named document retrieval theory which would then be part of library science. However, he commented that those working in the field of information retrieval were making the explicit claim to be working with information not documentation' [26, pp. 187–188].

What Brookes stated was:

> From an information science point of view, research on IR systems offers only a theoretical cul-de-sac. It leads nowhere. The anomaly I have noted is this: the information-handling processes of the computers used for IR systems, their storage capacities, their input and internal information transmissions, are measured in terms of *Shannon theory* measures – in bits, megabits per second, and so forth. On the other hand, in the theories of information retrieval effectiveness information is measured in what I call *physical* measures – that is, the documents (or document surrogates) are counted as relevant or non-relevant and simple ratios of these numbers are used. The subsequent probabilistic calculations are made as though the documents were physical things (as, of course, they are in part), yet the whole enterprise is called *information* retrieval theory. So why, I ask, are *logarithmic* measures of information used in the theory of the machine and *linear* or physical measures of information in IR theory?
>
> If *information* retrieval theory were called *document* retrieval theory, the anomaly would disappear. And document retrieval theory would fall into place as a component of *library* science, which is similarly concerned with documents. But that is too simple an idea. Those who work on IR theory explicitly claim to be working on *information*, not *documentation*. I therefore abandon the simple explanation of a misuse of terminology. I have to assume that IR theorists mean what they say – that they are contributing to *information science*. But are they? [27, p. 2; emphases in original].

In my view *it is not too simple an idea* that information retrieval theory is in reality document retrieval theory and thus closely associated with library science.[10] It is not difficult to disprove Brookes' statement that information retrieval is not dealing with documents. A short orientation in the literature demonstrates this, and even if the Cranfield experiments spoke about 'information retrieval', their modern continuation, the TREC experiments are speaking about 'text retrieval'. 'Text retrieval' and 'document retrieval' are often used as synonyms for IR. It is very embarrassing to us that these kinds of elementary misunderstandings are so widespread, and I find it extremely urgent to build a more solid theoretical basis for our field.

The view of knowledge as isolated facts or ideas is related to empiricism and rationalism, whereas pragmatic epistemology looks at knowledge as a collection of theories fulfilling some purpose for living organisms. Each individual is influenced by some presumptions, conceptions, and theoretical influences. This is true both for information scientists, the users, and all other people (in a broader understanding even for animals). Such influences can work through language and other cultural phenomena, and thus influence in a more or less collective way. Each individual has thus a certain theoretical makeup, which is shaped by some specific

---

[10]In science it is generally recognised that simple ideas should be preferred to complex ideas. This principle is known as 'Ockham's razor' after William of Ockham (1287–1347). The motive for information scientists not to want to be regarded as related to library science might be that important technological improvements were not done by people associated with librarianship, but with computer science.

contexts and from some assumptions. This makeup influences the ways in which all texts are written as well as how they are read. It also influences the way information systems and services are designed and managed. The production and evaluation of knowledge cannot be done by empiricist or rationalist principles alone, but by a combination, in addition to historical knowledge of the origin of the theories and in considering human goals and values. Knowledge becomes much more contextualised, which brings the documents and their content into focus. A document has a history, one or more authors or producers, a connection to other documents, and so on. All this is very well known and understood in many areas in the humanities, where there are disciplines such as the history of literature, criticism of documents (including films), and source criticism in history, but often less well understood in technological fields.

My conclusion of this section is that information is not a thing, but that *all* things can be informative [cf. 10, p. 50] – to a greater or lesser degree, and always only from the point of view of specific situations. Things that are generally seen as important because of their informative potentialities can be termed documents, and if they are judged collectively important, they are collected, organised, retrieved and disseminated by archives, libraries, museums, journals, databases, and other kinds of memory institutions.[11] Documents themselves (e.g. journals) are increasingly created in electronic forms, and so are memory institutions (e.g. digital libraries). Even if documents are electronic, they are still not information (but are potentially informative). Documents and memory institutions are information resources, and the aim of library science/documentation/IS is to facilitate the utilisation of such information resources. The kind of expertise needed for the management of memory institutions is not primarily IT expertise, but expertise on information/knowledge resources – which to a high degree is subject/task specific (e.g. resources relevant for high school education). The core functions provided by librarians/information specialists are related to such tasks as selection of documents, their indexing and classification, and searching/retrieval of information and documents for users. Some prominent researchers in computer science and IS, e.g. Salton [28], have seen this problem as purely technological, where there is no need for human expertise: librarians/information specialists and human interpreters are going to disappear in future information systems. This might turn out to be the case, but as information specialists (and educators of these) we must consider very carefully what kind of expertise is involved. To answer this question is to formulate a theory about indexing, seeking and retrieving documents. Salton's approach is just one among others, and even if it is very impressive, it is not necessarily the most true or useful perspective for IS, and it is quite different from the approach that I from a different platform find it necessary to develop [cf. 19].

4. DOCUMENTS AND MEMORY INSTITUTIONS

Figure 1 reveals an understanding of (L)IS identical with the one that I am endeavouring to formulate. It implicitly regards the processes in LIS in

---

[11]Private documents such as diaries are normally not of such collective interest that they are collected in memory institutions.

The Branches of Knowledge

 …

  Preservation of Knowledge

   Institutions and Techniques for the Collection, Storage, Dissemination, and Preservation of Knowledge

    Protection and storage of objects and artefacts

     Museums and galleries

     Libraries

     Historical places and landmarks

     Public and private collections of animals and plants

     Parks and nature preserves

    Storage and retrieval of information [informative objects]

     Dictionaries and lexicons

     Encyclopaedias

     Atlases and map collections

     Libraries

     Archives

     (Bibliographic and numeric databases)

     (Magnetic and optical recording)

     Almanacs

    Institutions for the advancement and dissemination of knowledge

     Educational institutions

     Academies of learning, or societies established for the advancement of knowledge

     Publishing; selection, preparation, and marketing of printed material

     Broadcasting

     Observatories and planetariums

Figure 1. *LIS classification in Encyclopædia Britannica* [23][12]

a social and cultural perspective. In the area of printed documents, a division of labour has evolved between, for example, libraries, archives, museums and other kinds of 'memory institutions', including journals and the systems of primary, secondary and tertiary literature, source literature and 'repackaged literature' (e.g. textbooks). In the area of electronic communication and the Internet, all these institutions are going to change and to use the same basic medium of communication, the divisions of roles have to be redefined, and many concepts such as 'document' also have to be redefined. The theoretical approaches in the study of those 'memory institutions' and their processes must be redefined as well. The important problem is to develop a theoretical approach to information resources, memory institutions, and to indexing, seeking and retrieving documents.

---

[12]The systematic arrangement in *Britannica* from which this classification is taken is called 'Spectrum' and is constructed by the editors. It is not identical with the structure of the article in *Britannica* written by the information scientist V. Slamecka: 'Information processing and information systems'.

Britannica [23] does not offer a single theoretical formulation of LIS. Partly its approach can be termed idiographic[13] because many articles are about concrete LIS institutions. Such an approach would probably not satisfy most researchers in IS, because we would like to discover or construct general principles for indexing, seeking and retrieving documents, even though idiographic research should be recognised as a very important supplementary approach. Another approach implicit in *Britannica* – and in most texts on IS – is eclecticism, the introduction of various, perhaps conflicting approaches. IT perspectives, cognitive perspectives, cultural perspectives etc. may be introduced in different contexts without attempting to develop an overall consistent view.

A more genuine approach to the study of all kinds of 'generalised texts'[14] is semiotics, which is an approach that some information scientists are also turning towards. It is also explicitly mentioned in the article in *Britannica* 'Information processing and information systems' [23]. Semiotics is defined as the study of signs, and signs are different from isolated facts or ideas in important respects. Firstly, they are material realities, not just mental phenomena. Secondly, according to the founder of semiotics, Charles Sanders Peirce, a sign represents something to somebody, it is thus not objective in the positivist sense, and semiotics has the potential to study how meaning is established culturally. Semiotics is seen by an increasing number of researchers as an alternative to the methodological individualism of the traditional cognitive sciences. Space does not allow a deeper analysis of the semiotic approach to IS here. An influential introduction to communication studies by Fiske [29] identified two major paradigms in communication studies: (1) the process-analytic paradigm founded by Shannon and Weaver [20]; and (2) the semiotic paradigm founded by C.S. Peirce and Ferdinand de Saussure. The differences between these two approaches may indicate a kind of paradigm shift in communication studies, which may turn out to be a forerunner of a similar paradigm shift in IS. Information scientists have mostly used the process-analytical tradition – without realising that alternative traditions do exist. This has had serious consequences for the kind of research done, and especially, what research has not been done in IS.

Semiotics is, however, only one among other approaches, which tries to form alternatives to positivistic approaches. Other alternatives are hermeneutics, social constructivism, grounded theory, feminist epistemology, activity theory, etc. I agree with Vickery [30, p. 458] that contributions from such approaches have hitherto often lacked the necessary degree of specificity concerning problems in IS. Nonetheless, I find a reorientation of the field towards socially, culturally and historically oriented approaches mandatory [cf. 31], and I can refer to some of my own contributions [32, 33] as claims of specific contributions to IS informed by such an approach.

---

[13]The formulation of 'idiographic' as opposed to 'nomothetic' approaches was done by W. Windelband (1848–1915).

[14]The concept of 'text' in semiotics is very broad, including, for example, pictures. A text is something one learns to read (in a very general sense) in a culture (or in a science). See Hjerppe [11]. This conception of text is closely related to the general concept of 'document' in IS.

The conclusion of this section is that the object of LIS should be seen as cultural institutions (in particular science, scientific documents, (sub)languages, scientific communication and electronic databases understood as cultural institutions). From a LIS perspective it is important to consider memory institutions and documents from a user's perspective; to develop norms of efficiency and knowledge on how to optimise these institutions (which is not the same as their automation, even if automation is a very efficient way to improve their performance). The criteria and norms for the optimal management of these institutions are not primarily technological or individualist-cognitivist criteria, but criteria for the advancement of knowledge, building on collective experiences. We are back to problems regarding relevance, meaning, and the role of different kinds of 'texts', documents and systems in processes of learning and discovery; we are back to 'cognitive' kinds of problems.

Information seekers' behaviour should be studied in their rich environment, related to all kinds of documents, memory institutions, knowledge areas, problem situations, etc. Information seeking is not just reduced to fact retrieval, to the matching of terms, to input and output, or similar matters. The information seeker is understood as a person in a specific cultural context and the person's information needs as well as the information resources and the memory institutions are all influenced by the same kind of meaning-producing forces (e.g. language and metatheoretical trends). Such a non-reductionist approach should not be mistaken as a kind of anti-scientific approach, which opposes the formulation of general principles.

Information specialists, who are seeking information, selecting and organising documents etc., have to deal with different subject areas, and the necessity to have adequate subject knowledge should not be underestimated. A central element in a theory of information seeking, indexing etc. must be a theory about the information seekers' and indexers' interpretation of the sources and concepts in the field. How can this be generalised from different subject domains?

The most general knowledge we have about different interpretations of knowledge fields is the metatheoretical 'paradigms' and philosophical approaches to the field. Therefore epistemology should be regarded a core discipline in IS. This can be seen as an agreement with the basic claim in the cognitive view: 'that any processing of information, whether perceptual or symbolic, is mediated by a system of categories or concepts which, for the information processing device, are a model of his [its] world' [34, pp. xiv–xvii; cf. 22, p. 16]. Epistemological theories are our most general models of how people look at their respective fields. Researchers who identify themselves with the cognitive view have not yet drawn this conclusion. Even though epistemology is mentioned in the literature [e.g. 22, p. 8, fig. 1.2] its real importance has not so far been acknowledged.

8. CONCLUSION

The problem that I have raised in this article is whether we should prefer the term 'documentation science' (as recently introduced in Tromsø) or 'information science' (as recently introduced in Copenhagen). I have tried to argue that the conceptions of information, information retrieval and information

science[15] are seriously flawed, and that the problems in IS are not just termi-nological but rooted in problematic theoretical assumptions. This blocking in our field can to a large degree be avoided by changing the object of study from mental phenomena of ideas, facts and opinion, to social phenomena of com-munication, documents and memory institutions. This is a strong argument for choosing the expression 'documentation science'. This view is also supported by White and McCain [37, p. 353] who suggested that information science is really about literatures, and that much use of the term 'information' in our field is misdirected.

The article could stop here. Tromsø won. However, the terms LIS and IS are rather well established, and they can in my opinion be justified *if* we make it clear that we are studying potentially informative objects. The most important thing to realise is that the intrinsic natures of these objects are relatively irrelevant. It is their informative functions which are of primary interest to us. This calls for approaches in IS which consider the social contexts in which their meanings and the needs for them are produced. The sociocognitive approach [38] is an impor-tant candidate for such a reorientation.

REFERENCES

1. Schrettinger, M. *Versuch eines vollständigen Lehrbuchs der Bibliothekswis-senschaft.* Band 1–2. München, 1808–1829.
2. Kunze, H. and Rückl, G. *Lexikon des Bibliothekswesens.* Band 1. Leipzig: VEB Verlag für Buch und Bibliothekswesen, 1974.
3. Butler, P. *An introduction to library science.* Chicago: University of Chicago Press, 1933.
4. Ranganathan, S.R. *Preface to library science.* Delhi: University of Delhi, 1948.
5. Ranganathan, S.R. *The five laws of library science.* 2nd ed. London: Blunt and Sons, 1957.
6. Olaisen, J. Alternative paradigms in library science: the case for para-digmatic tolerance and pluralism. *Libri*, 35(2), 1985, 129–150.
7. Rayward, W.B. Visions of Xanadu: Paul Otlet (1868–1944) and hypertext. *Journal of the American Society for Information Science*, 45(4), 1994, 235–250.

---

[15]Fields outside library and documentation research can also claim to be information sciences [cf. 35, 36]. In Denmark the University of Aarhus has one programme in infor-mation science (not related to LIS), and the Royal School of Library and Information Science another (both described in Danish in the article 'informationsvidenskab' in the *National Encyclopaedia* by Peter Bøgh Andersen and Peter Ingwersen, respectively).

8. Otlet, P. *Traité de documentation. Le livre sur le livre. Théorie et pratique.* Bruxelles: Editions Mundaneum, 1934.

9. Meadows, A.J. Theory in information science. *Journal of Information Science*, 16, 1990, 59–63.

10. Buckland, M.K. *Information and information systems.* New York: Greenwood, 1991.

11. Hjerppe, R. A framework for the description of generalized documents. *Advances in Knowledge Organization*, 4, 1994, 173–180.

12. Griffith, B.C., ed. *Key papers in information science.* New York: Knowledge Industry Publications, 1980.

13. Borko, H. Information science: what is it? *American Documentation*, 19(1), 1968, 3–5.

14. Schrader, A. M. *Toward a theory of library and information science.* Vol. 1–2. Ann Arbor: University Microfilms International, 1983.

15. Hahn, T.B. Text retrieval online: historical perspective on web search engines. *ASIS Bulletin*, Apr/May 1998. Also available at: http://www.asis.org/Bulletin/Apr-98/hahn.html (visited 22 January 1999).

16. van Rijsbergen, C.J. A new theoretical framework for information retrieval. In: Association for Computing Machinery. *Proceedings of the 1986 ACM Conference on research and development in information retrieval.* New York: ACM Press, 1986, 194–200.

17. Betchel, W. and Graham, G., eds. *A companion to cognitive science.* Oxford: Blackwell Publishers, 1998.

18. Hjørland, B. and Albrechtsen, H. Toward a new horizon in information science: domain analysis. *Journal of the American Society for Information Science*, 46(6), 1995, 400–425.

19. Hjørland, B. *Information seeking and subject representation. An activity-theoretical approach to information science.* Westport, London: Greenwood Press, 1997.

20. Shannon, C. and Weaver, W. *The mathematical theory of communication.* Illinois: University of Illinois Press, 1949.

21. Belkin, N. *A concept of information for information science.* London: University of London, 1977. (Unpublished doctoral dissertation)

22. Ingwersen, P. *Information retrieval interaction.* London: Taylor Graham, 1992.

23. Slamecka, V. Information processing and information systems. In: *Britannica® CD 99 Multimedia Edition.* Chicago, Il.: Encyclopædia Britannica, Inc., ©1994–1999.

24. Sparck Jones, K. Architecture problems in the construction of expert systems for document retrieval. In: Wormell, I., ed. *Knowledge engineering: expert systems and information retrieval.* London: Taylor Graham, 1987, 7–33.

25. Spang-Hanssen, H. *How to teach about information as related to documentation.* Unpublished manuscript, 1970.

26. Ellis, D. *Progress and problems in information retrieval.* London: Library Association, 1996.

27. Brookes, B.C. Information technology and the science of information. In: Oddy, R.N. et al., eds. *Information retrieval research.* London: Butterworths, 1981, 1–8.

28. Salton, G. Letter to the editor. *Journal of the American Society for Information Science*, 47(4), 1996, 333.

29. Fiske, J. *Introduction to communication studies.* 2nd ed. London: Routledge, 1990.

30. Vickery, B. Metatheory and information science. *Journal of Documentation*, 53(5), 1997, 457–476.

31. Hjørland, B. Theory and metatheory of information science: a new interpretation. *Journal of Documentation*, 54(5), 1998, 606–621.

32. Hjørland, B. Information retrieval, text composition, and semantics. *Knowledge Organization*, 25(1/2), 1998, 16–31.

33. Hjørland, B. The classification of psychology: a case study in the classification of a knowledge field. *Knowledge Organization*, 24(4), 1998, 162–201.

34. de May, M. The cognitive viewpoint: its development and its scope. In: *CC 77: International Workshop on the Cognitive Viewpoint*. Ghent: Ghent University, 1977, xiv–xxxii.

35. Hjørland, B. Review of: Stonier, T. *Information and meaning: an evolutionary perspective*. Berlin: Springer, 1997. *World Futures: the Journal of General Evolution*, 53(4), 1999, 371–375.

36. Marijuan, P.C. First conference on foundations of information science: from computers and quantum physics to cells, nervous systems, and societies. *Biosystems*, 38, 1996, 87–96.

37. White, H.D. and McCain, K.W. Visualizing a discipline: an author co-citation analysis of information science, 1972–1995. *Journal of the American Society for Information Science*, 49(4), 1998, 327–355.

38. Jacob, E.K. and Shaw, D. Sociocognitive perspectives on representation. In: Williams, M.E., ed. *Annual Review of Information Science and Technology*, 33, 1998, 131–185. Medford, NJ: Information Today for American Society for Information Science.

(*Revised version received 4 May 1999*)