

The phenomena of interest to Information Science*

Gernot Wersig & Ulrich Neveling

Discusses the various explicit and implicit definitions of information and information science, against a view of their historical development. Shows how the various views of information science overlap with other disciplines, and concludes with a proposal for a definition of information science based on social need. A schema of information sciences is put forward with the plea that any discussion of information and information science should first declare the definitions to be used.

1. Different origins of information science

"Information science", "informatics" and whatever else it is called is a newly emerged field of study whose first consciousness of being a discipline dates back perhaps to the late fifties (although scientific work had been undertaken before by individual researchers). This is one of the very few statements that could be made concerning "information science" which would not be discussed at length by a set of randomly collected people calling themselves "information scientists".

"Information science" did not develop out of another field of study (like psychology) or from the intersection of two fields (like bio-chemistry), but out of the needs of an area of practical work, called 'documentation' or 'information retrieval'. Although the introduction of new technologies, particularly electronic data processing, made the emergence of this discipline necessary, the contributions to the birth of "information science" came from many different

* This paper was published in the *Information Scientist*, vol.9 no.4, pp 127-140, December 1975.

disciplines (due to the different backgrounds of people entering a field where no established educational system existed) and were caused by a lot of different interests (due to the different areas of application involved in information work). Some of them are

- computer science (as the technology played an important role)
- library science (as a lot of people working in the field were trained librarians)
- philosophy and taxonomy (as the phenomena of classification played an important role)
- linguistics (as natural language played an important role as both the object and the instrument of practical work)
- information theory (perhaps due to the terminological similarity)
- cybernetics (since everybody at that time tried to play with cybernetic models)
- mathematics (since there are always people trying to apply reality to some beautiful mathematical model).

It is interesting to note that disciplines like social sciences, science of science and communication theory entered rather late into the discussion about "information science".

This variety of approaches led to a situation where every participant in the discussion might agree that there should be something like "information science", but that it had to be his information science relying on his special background.

The difficulties in discussing the phenomena which may belong to "information science" are partly caused by

- the different backgrounds of the participants (see above).
- the missing historical derivation of the whole field.

Since information science emerged at a specific time the historical processes which led to the emergence of information science should be studied. They could provide a more general understanding of what we are talking about.

- the terminological differences in the use of the term 'information'. This problem will be discussed in section 2. We face here a situation that many different approaches are possible to fill the term 'information' with semantic content; that all possibilities are used but usually not expressed; but that there is no generally accepted criterion which can determine the choice.

At this point a personal view may be expressed. Since in our opinion science is not something that can be justified in itself but is always justified by some social needs which are to be served by this science, we have to find out what the specific reasons were that led to the development of this science. The main questions to which the answers may perhaps create some common understanding are:

- Why did information science emerge?
- Which social needs are to be fulfilled by information science?

In fact, it seems that more views have been expressed on information science than people exist who have been expressing views on this topic. A first attempt to systematize these views could perhaps be to classify them broadly into the following main categories:

- the phenomenon-oriented view
- the means-oriented view
- the technological-oriented view
- the purpose-oriented view

a) Phenomenon-oriented view

This category implies that there are a lot of people

stating that there is a phenomenon called 'information' and that "information science" is the science concerned with this phenomenon 'information'. In fact, this seems to be the largest group of people and - because of the difficulty with the term 'information' - the group with the most divergent views on the topic. The different possible definitions of the basic term 'information' will be given in section 2.

b) Means-oriented view

Some people, usually those coming from practice-oriented areas, try to define "information science" from their view of its application and therefore concentrate on the means which have to be used in practical information work. This is the case with people concentrating on classification theory (eg the Ranganathan school) or with people concentrating on the design of computerized information systems (eg Kunz/Rittel¹) or with people coming from librarianship.

c) Technological-oriented view

"Information science" is often seen only as that subset of computer science concerned with the application of electronic data processing equipment to the storage and retrieval of data. This, of course, is the case with people from computer science.

d) Purpose-oriented view

A few people, mostly those coming from the social sciences, define "information science" from the point of view that there are some social needs which have to be fulfilled and that 'information science' should serve and develop the practical work related to those needs.

2. 'Information' as the possible object of information science

The most extreme case of polysemy in the technical communication of information and documentation is the term

'information'. A semantic analysis made by one of the authors showed that there are at least six different approaches to the use and understanding of the term in the whole field of disciplines. Obviously each use and understanding of the term is justified but as ambiguity is one major fault in scientific communication and theory building, a rule has to be found to assess which understanding may be appropriate for which purpose.

Six major types or approaches to 'information' may be characterized based on the general structure of relations between humans and the world.²

a) Structure approach (matter-oriented)

The matter-oriented approach, usually leading to something that may be called STRUCTURE-definition, states that the structures of the world which may be perceived or not are 'information'. 'Information' is thus independent of whether a human being gathers it or not. There are several variants of this approach:

- information is given by the static relations between material objects: "We would define information as the position of all the atoms and molecules in the universe and of all sets and combinations of those atoms and molecules at any time".³
- information are the relations which may become perceptible if changes in the states of the physical objects occur.⁴
- information is a characteristic of physical objects.⁵

It is interesting to note that this attitude is mostly assumed by philosophers and is used both by idealists and Marxists.

b) Knowledge approach

The KNOWLEDGE approach states that the knowledge built up on the basis of perception of the structure of the world is 'information'. To us this seems to be a very dangerous approach because 'knowledge' itself is a polysemic term and sometimes it seems that people who do not know how to understand 'knowledge' replace this term by 'information' and vice versa.

Again there are some variants:

- the 'knowledge' is objectively given and may be actualized or not.⁶
- the 'knowledge' has to be acquired by at least one subject: "Information is knowledge concerned principally with qualitative concepts or ideas".⁷
- the 'knowledge' serves a specific purpose: "Information is data of value in decision-making".⁸
- the 'knowledge' must not only be acquired by available in communicable forms: "Information is knowledge being communicated".⁹

Here it is interesting to note that the approach is spread widely throughout the disciplines (perhaps because of the ambiguity of the term 'knowledge'); but that it is most acceptable to people concerned with decision theory (where a decision is often seen as the process of bridging gaps of 'knowledge' or 'information').

c) Message approach

Due to mathematical communication theory, 'information' is often equated with 'message' (as the information content is measured by the message). Again some variants can be identified:

- Message as a physical process:
 "....I mean the physical substance, in the form of energy, of the idea or concept which is composed from

one mind to another, ultimately by every transference".¹⁰

- Message as a set of symbols:
 "....whereby information....is defined as symbols being produced by a communicator to realize his communicative intent" (translated from 11).
- Message as a unit of physical carrier and symbols arranged on it:
 "Information is a unit consisting of semantic (something conceptually understandable) and a physical carrier" (translated from 12).

Obviously this approach is used mostly by people concerned with mathematical communication theory.

d) Meaning approach (characteristic of message-oriented)

Whereas the MESSAGE approach widely neglects the meaning side of the message, another type of definition, the MEANING-definition, accepts only the meaning of the message as 'information'.

Typical of this approach is an American standard: "The meaning assigned to data by known convention".¹³ As 'meaning' is at least as ambiguous as 'information' one may again suppose that people who do not know what 'information' is call it by another name which is as unclear as the first. Mostly, this is done by data processing people who perhaps understanding meaning better than linguists. On the other hand, linguists do the same only in the reverse way: "The meaning of a sign is the information about the designated object embodied, fixed and expressed by the sign". (translated from 14).

It would be interesting to listen to a discussion between linguists and computer scientists about the question "which was first - meaning or information?"

e) Effect approach (recipient-oriented)

According to the general structure of communication (and nobody doubts that 'information' has something to do with communication) one group of people places 'information' with the recipient (unlike the previous approaches which place it with the communicator) stating that 'information' occurs only as a specific effect of a specific (or unspecified) process. Again some variants may be found:

- 'Information' is the result of an unspecified process
 "'Information' is data produced as a result of a process upon data. That process may simply be one of transmission it may be one of selection; it may be one of organization; it may be one of analysis".¹⁵
- 'Information' is abstraction of knowledge
 "In fact, we say that information is that which changes what we know...." ¹⁶
- 'Information is reduction of uncertainty
 "A communicative action contains.....information, if and only if by this action uncertainty is removed or reduced for the perceiver" (translated from 17)

It is interesting to note that the last two variants are mostly used by behavioural scientists and that the last variant is - although rather psychologically biased - in accordance with the theoretical model of mathematical communication theory (in which the amount of information carried by a symbol or group of symbols relates to the probability of occurrence of that symbol or group).

(f) Process approach

Some people see 'information' not as one of the components of processes but as a process itself. This may be

- a process of human data processing
 "Information....is a process which occurs within a human mind when a problem and data useful for its solution are brought into productive union".¹⁸

- a communication process with a specific purpose
 "Information...is a set of purpose-directed actions covering the following elements: generation...., transmission, storage and enquiry". (translated from 19)

This diversity of views on 'information' is frightening, and in fact is probably incomplete. It would be a relief if we could follow Fairthorne's proposal: "Clearly, 'information' and its derivatives are words to avoid",²⁰ but obviously this is not possible. Many views exist only because of the interest in this phenomenon, and at least this term helps to promote ideas which would otherwise need far more effort to be conveyed to the right people. Who of us would like to follow an earlier proposal to call information science "Ichneutics"²¹ coming from the Greek "Ichneumon" (an Egyptian cat)? If the term 'information', or one of its derivatives like 'informatics', is unavoidable, we should make it clear in every instance what is meant.

3. Proposals for the interpretation of 'information' in "information science"

3.1 Societal responsibility

I suppose we all call ourselves "information scientists". A young English lady one of the authors met last year who had taken a course of several weeks and is now working as an information officer, called herself an "information scientist". Should and could we do something to counter this misuse? If not, it does not matter, because we could all be called "information scientists" by other people based on the simple fact that we all take part in a meeting concerned with "information science". Our hypothesis would be that on a first view nothing more connects us than a common characteristic of being at the same meeting or being called by the same name. But we do not think that this is true. Of course, we all have divergent views on the theoretical foundation of what we call "information science" and this is perhaps a situation where this diversity could be creative.

On the contrary, we think that there have been social developments which have created comparable languages, discussing common problems, in common terms. It is only possible now to outline these developments briefly.

Perhaps the problem started when Gutenberg invented a new communication technology which made it possible to transfer pieces of knowledge to a larger public. In the 17th and 18th century it was felt that original publications of scientific value could not be distributed to the interested community (scientific and non-scientific) so secondary publications emerged (in fact the first technical journals were abstract journals; the movement of the encyclopaedists could be considered under the same label; later on 'Moralische Wochenschriften' and 'Intelligenzblätter' in Germany and their equivalents in other European countries followed similar lines).

With the increasing relevance of science for industrial development in the 19th century (first industrial revolution) specialization of science began, more sciences were produced and needed, and scientific communication grew and grew. This process is still under way. Briefly stated, this led to a situation where an 'anonymous' group emerged between the producers of data, findings and discoveries on the one hand and those people needing the data, findings and discoveries on the other. Consequently, in the process of the division of labour a new field of practical activity emerged and an older profession started to change its definition. The information people appeared - with different titles in different countries, but similar in function. And this function is - again described very simply - to ensure that those people needing some piece of knowledge for their scientific and technical work (not now restricted to science and technology but including 'occupational roles' or roles even less restricted) will get

it or not. This is not a widespread philosophy but a social function derivable from historical development. At first they tried to solve their problem by doing practical work but as

- information needs changed and grew
- the social communication processes became more and more complicated
- scientific and technical communication reached a qualitatively new stage by the 'scientific-technical revolution' and a worldwide dimension
- new and complicated communication and information technologies developed,

the purely practical approach with traditional, mostly library, methods proved to be inefficient in solving the underlying problem. Out of the needs of a practice which grew and became more and more complicated, scientific work emerged and by and by a set of people appeared, a new terminology was used, and specialized communication took place. Thus a new discipline developed - not because of a specific phenomenon which had always existed and which now became an object of study - but because of a new necessity to study a problem which had completely changed its relevance for society. Nowadays the problem of transmitting knowledge to those who need it is a social responsibility, and this social responsibility seems to be the real background of "information science".

By this 'credo' we do not deny the possibility of working scientifically for pleasure, or of pursuing information science in the search for truth. But the individual feelings of information scientists do not add up to a discipline which society should be expected to support.

3.2 The place of the discipline

The development of an "information science" which has to attempt to overcome some of the effects of specialization in science and technology, is itself a

phenomenon of the differentiation of disciplines. Therefore one of the main questions we have to face is how this new discipline fits into the century-old web of disciplines. In looking at this problem two different answers have to be given.

First: Gathering, arranging, rearranging, and evaluating pieces of knowledge is an integral part of every scientific or technical discipline. Therefore, regardless of how the discipline under consideration is defined, every discipline concerned with methods and techniques of handling these pieces of knowledge will serve as an auxiliary discipline for every other discipline (like mathematics as a discipline providing tools for a specific kind of representation and transformation of messages).

Second: Unlike mathematics, information science is based on a specific social need and its object - again regardless of how it is defined - has to include these social processes. Therefore information science is a discipline which has to define its boundaries vis-a-vis other object-oriented disciplines. Many proposals have been made, to which we will not add, but we will present some alternatives from which the proposal should be chosen which best allows the fulfilment of the social needs underlying the development of this discipline. Other alternatives may be possible but those presented here are based on an analysis of the practicability of different proposals:

a) The broad solution

The broad solution may be based on the STRUCTURE approach to information. If every structure of the objective world is 'information' then a science concerned with the methods of discovering this information, of representing it, and of transforming it into other representations which permit further conclusions, will be possible and helpful for all kinds of scientific work.

Such a science will then include

- part of mathematics
- logic
- philosophy of science
- philosophy of language
- transformational grammar
- mathematical communication theory

Such a discipline will, of course, be a metascience, but will not solve the problems which led originally to the development of information science as far as they are contained in this paper.

Questions particularly raised by using this approach may be

- what are objective structures of STRUCTURE, ie of reality
- what is the nature of STRUCTURE, ie of reality
- how can the STRUCTURES be discovered
- what is the nature of scientific discovery
- what is the nature of the relation between mind and reality
- how may discoveries be described in an adequate way, ie how is reality as reflected by the human mind objectively describable
- how may different forms of description of reality be interpreted in an adequate way (hermeneutics)
- how may different forms of description be used to gain new insights or discoveries (secondary discoveries).

Obviously, the main relation underlying this approach would be the relation between man as the discoverer and reality. This would lead unavoidably to a more philosophical "information science", being to some extent a metascience for all other sciences. It is interesting to note that this solution is only phenomenon-oriented

(according to our previous classification) as all means, technologies and purposes are allowed as objects of study.

b) The medium solution

The medium solution may be based on the combination of the KNOWLEDGE and the MESSAGE approach. It could be described as follows_ The basic notion is 'information' as recorded knowledge. Then information science will be concerned with studying

- knowledge structures
- the relation between knowledge structures and modes of recording them
- the methods, means and structures of recording knowledge elements
- the methods, means and effects of transferring recorded statements
- the methods of interpreting recorded statements for their knowledge content (again a variant of hermeneutics)
- the effects that recorded statements of knowledge can have on those receiving and decoding them
- the relations between recorded statement of knowledge and behaviour
- the purposes for which recorded statements of knowledge may be produced, transmitted and used.

At a first glance this medium solution is only phenomenon-oriented but may produce variants by combination with

- the technological-oriented approach (eg if the phenomenon is looked at only from the point of view of processing recorded statements of knowledge by EDP as in some areas of applied information science)
- the purpose-oriented approach (eg if only those recorded statements of knowledge are considered

which are used for a specific purpose in decision-making²² or in transforming a state of knowledge²³).

The following disciplines will contribute to such an "information science":

- psychology, particularly cognitive psychology and psycholinguistics
- mathematical information theory, particularly coding theory
- communication theory
- computer science
- linguistics, especially semantics and pragmatics

c) The narrow solution

The broad and medium solutions were similar in having no limitation of the area in which information is considered but did differ in the generality of the basic concept of information. The narrow solution (as it is proposed here) does not give a narrow understanding of 'information' but of the area in which 'information' is considered. This is, of course, the solution which is not (like the broad and the medium solution) based on abstract, theoretical interests but on practical problems which have to be solved.

There are different possible limitations of the area in which information problems are considered. The traditional limitation of science and technology is the one mostly used (eg Michailov/Cernyi/Giljarevsky²⁴) and which has recently been criticized as being too narrow. This is an approach which is rather phenomenon-oriented as only the phenomenon of information, usually defined as recorded knowledge in the context of scientific and technological work, is considered. But there is some combination with the purpose-oriented approach because scientific and technological information is not studied for its content but to see how it might be collected, organized and distributed.

A way of limiting the area more broadly may be offered by the following sequence_

there are scientists and technicians -
 who produce data (recorded knowledge -
 which are needed by themselves -
 but need other kinds of data (not
 produced by scientists or technicians).

We could state that information science in this sense has to see that

- scientists and technicians get all data (scientific/technological or not)
- scientific or technological produced data are distributed to all people needing them.

This would equate with the purpose-oriented approach. Even this limitation seems to be too narrow though a purpose-oriented approach needs some limitation in order to serve a practical purpose. A proposal which could be made, allowing a practical purpose approach while being still broad enough for practical applications (eg information systems in public administration) would have the following constraints:

- a distinction is made between organized and non-organized social communication, and the area of interest restricted to organized communication.
- a distinction is made between socially and individually justifiable information needs, and the area of interest restricted to those which are socially justifiable.
- a distinction is made between information needs related to social service and those which are not, the area of interest being restricted to those related to social service.

It should be obvious to the recipients of this paper that the authors are very much in favour of the narrow solution as it seems to be the only one that allows a solution to the concrete problem we have to face. Of course, it would be possible to call this science by a name other than "information science", but as this science we are talking about is already connected with the term "information science" we would prefer to go on calling it "information science", particularly in view of the broad and medium solutions which are nothing other than fields of study already.

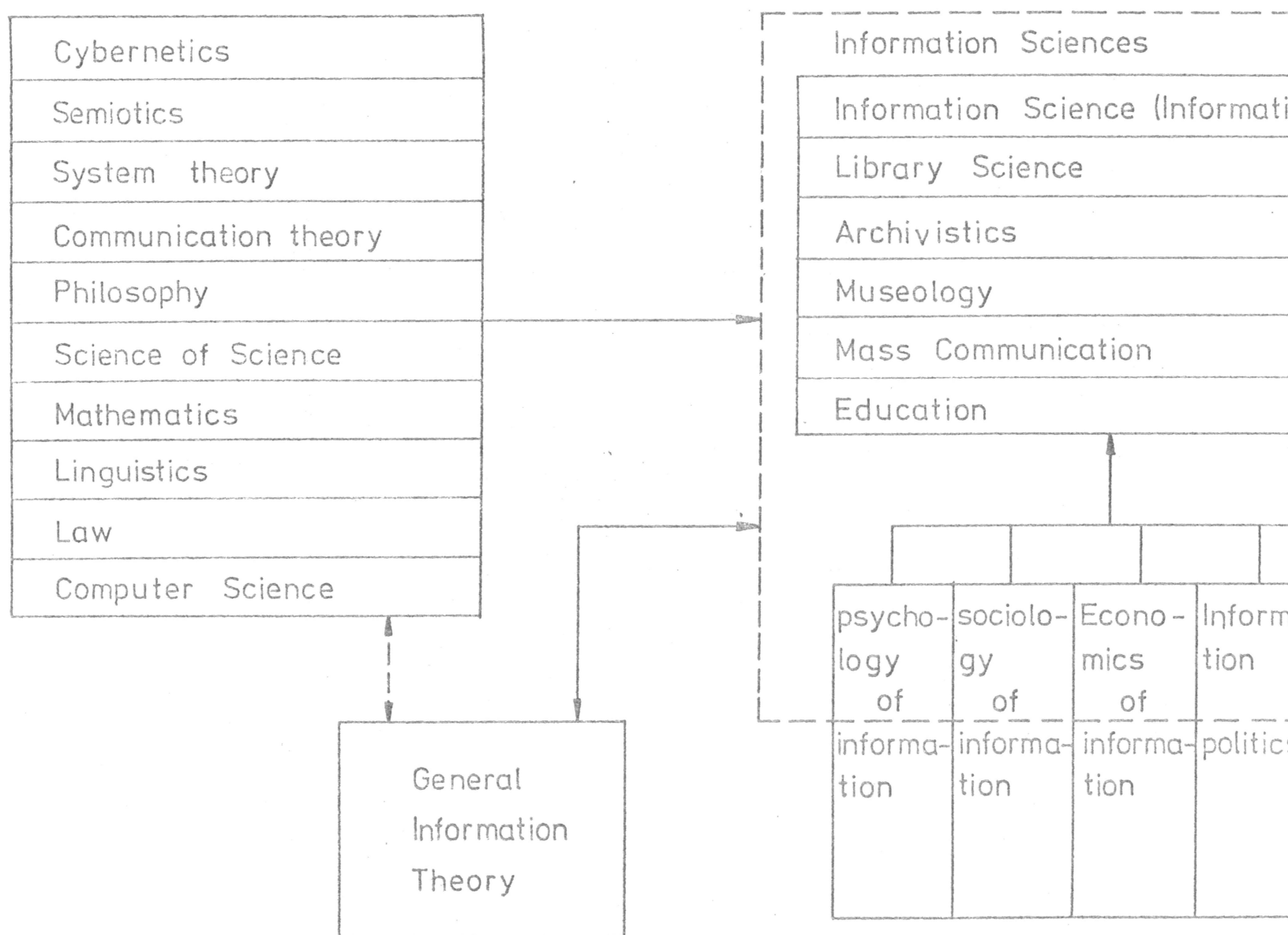
3.3 A proposal²⁵

From a combination of historical evolution, the development of specific societal needs and the development of new methodologies and technologies a new discipline has emerged which is sometimes called "information science" (or by other derivatives of the term 'information'). This science is based on the notion of the information needs of certain people involved in social labour, and of concern with the study of methods of organization of communication processes in a way which meets these information needs. The basic term 'information' can be understood only if it is defined in relation to these information needs

- either as reduction of uncertainty caused by communicated data
- or as data used for reducing uncertainty

This science is concerned with the organization of communication processes aimed at information for a specific clientèle. This is a science somewhat similar to mass communication aimed at the fulfilment of the socially and individually justifiable information needs of a large public. After this, similar disciplines are library science, museology, archivistics, education (all

FIGURE1 : The system of information sciences



catering for different clientèles according to different information needs).

On the other hand there are a lot of intersections of these disciplines with other traditional disciplines, eg

- psychology (psychology of information)
- sociology (sociology of information)
- economics (economy of information)
- political science (politics of information)
- technology (information technology)

All these disciplines concentrate on the study of information processes (ie communication processes aimed at the reduction of uncertainty) or components of them. Therefore it is possible to call all of them "information science" (the whole system of "information sciences" is given in Figure 1 overleaf).

There is a common subset of these disciplines and some other disciplines (eg those which would contribute to the broad or medium solution) which contribute to the information sciences. This subset is the whole set of efforts made to study 'information' in all of its facets (including all approaches specified in section 2). This field of scientific activity to which a lot of disciplines contribute could then be understood as 'General Information Theory' (of which statistical information theory forms a specialization). In all theoretical discussions on "information science" it should therefore clearly be stated whether the speaker or writer is discussing the phenomena, characteristics or methods of

- information science
- the information sciences, or
- General Information Theory

References

1. W. Kunz, H. Rittel: Die Informationswissenschaften. München-Wien (1972).
2. see G. Wersig: Information Kommunikation Dokumentation. Pullach. (1974).
3. E.G. Montgomery: Four 'new' sciences. AGARD. (1968).
4. V.A. Polushkin: On the determination of the concept 'Information' (Origl. Russ.) *Nauckno-Techn. Inform.* No. 9, p.6-8. (1963).
5. eg H. Vogel: Materie und 'Information'. *Dt.Zs.Philos.* Vol. 11, p. 314-323. (1963).
6. T. Kinder: The mathematical foundation of bibliographic organization. *Amer.Doc.* Vol.11, p.2-8. (1960).
7. B.E. Holm: How to manage your information. New York-Amsterdam-London. (1968).
8. M.C. Yovits: Information science: Towards the development of a true scientific discipline. *Amer.Doc.* Vol.12 p.191-197. (1961)
9. J.C. Costello Jr: Some solutions to operational problems in concept co-ordination. *Amer.Doc.* Vol.12, p.191-197. (1961).
10. E.E. Graziaur: On a theory of documentation. *Amer.Doc.* Vol. 19, p.85-89. (1968)
11. K. Koszyk, K.H. Pruys: dtv-Wörterbuch zur Publizistik. München. (1969).
12. G. Jedziny: *Kybernetik und Informationssysteme*. ZIID-Zs. Vol.15, No.3, p.113-116. (1968)

13. USA Standard vocabulary for information processing.
USASX3 12-1966. New York. (1966).
14. L.O. Resmikov: Erkenntnistheoretische Fragen der
Semiotik. Berlin. (1968)
15. R.M. Hayes: Education in information science. *Amer.
Doc.* Vol.20, p.362-365. (1969)
16. M.E. Maron: Mechanized documentation: The logic
behind a probabilistic interpretation: In: M.E. Stevens,
V.E. Giuliano, L.B. Heilprin (eds). Statistical associ-
ation methods for mechanized documentation. Washington
p.9-13. (1965)
17. H. Hörmann: Psychologie der Sprache. Berlin.
Heidelberg-New York. (1967)
18. A.G. Hoskovsky, R.J. Massey: Information science: Its
end, means and operations. In: Information transfer.
Proc.ASIS, Vol.5, New York. (1968).
19. J. Koblitz: *Information und Dokumentation* - als
Teilgebiet der wissenschaftlichen Information. ZIID-Zs.
Vol. 16, No. 1, p.36-43. (1968)
20. R.A. Fairthorne: 'Use' and 'mention' in the information
sciences. In: L.B. Heilprin, B.A. Markuson, L.Goodman
(eds): Proc. symposium for information science. Warrenton,
Va., Sept. 7-10, 1965. Washington. (1965)
21. H. Arntz
22. M.C. Yovits: A theoretical framework for the development
of information science. Paper presented at the FID/RI
meeting 1974 in Moscow.
23. N.J. Belkin: The concept of information in informatics.
Paper presented at the FID/RI meeting 1974 in Moscow.

24. A.I. Mikhailov, A.I. Chernyi, R.S. Giljarevskij:
Osnovy informatiki. Köln-Opladen. (1970)
25. G. Wersig: Informationssoziologie. Frankfurt. (1973)