Paper presented at Drexel Institute of Technology, Lecture Series in Information Science, Philadelphia, Pa,, September 29, 1964.

Is Science Information or Is Information Science?*

When Claire Schultz asked me to open up this series of lectures, I seem to recall her desire that I try to give you some idea of where I think the field of information science is going. I must confess that the title of my talk was created partly to be cute. However, the very ambiguity of the words in my title epitomizes many of the problems we face in information science. For this reason I felt it would be better to stick with the original title. To modify it for something less cute, but perhaps more pertinent, like "Whither Information Science?", would have been less provocative.

The vagaries of English are well exemplified in the permutations and combinations of the words science and information. The words taken separately have a significance quite distinct from their combined forms. Science is knowledge. Knowledge involves information. But the compound word science-information is another animal as is the word or noun phrase scientific information. And information-science is quite another animal. I happen to feel that information-science is the right choice for designating the science of information, whereas science-information concerns the information of science. This seemingly trivial observation becomes more significant when some of you recall that Mr. Leslie Wilson, Executive Director of Aslib, wrote in the pages of Special Libraries recently (1) that, in England at least, an Information Scientist is what we in the United States call a science-information specialist. Indeed,

*Paper presented at Drexel Institute of Technology, Lecture Series in Information Science, Philadelphia, Pa., September 29, 1964.

the Institute of Information Scientists, of which I am a member, Mr. Wilson would have us understand, is really an institute or society of science-information specialists.

Now we have in the United States an organization which has several hundred members who also call themselves information scientists. I include myself among these information scientists. At present, I am a member of the governing council of that organization. Some of you might have guessed that this is the American Society of Information Scientists. However, life is never quite that logical or simple. The organization in question is the American Documentation Institute. As soon as my Council term expires in a few months, I shall continue periodically to suggest that its name be changed to the American Society of Information Scientists.

We also have in the United States a large Documentation Division in the Special Libraries Association, of which I am also a member. We also have, in the Association of Computing Machinery, a Special Interest Group on Information Retrieval, of which I am also a member. The ACM people, especially SIGIR, are interested in the same problems as ADI and SLA. The list must also include the American Chemical Society Division of Chemical Literature, Medical Library Association, Association of Machine Translation and Computational Linguistics, and the new Classification Society. I am not listing these organizations to display my abilities as a joiner.

Rather, it is a somewhat painful way of getting across the point that much semantic confusion exists in the information science field. That confusion is not helped by those who feel it is an unfortunate accident of history; that the term information theory has been improperly usurped by a rather narrow group of probability theoreticians. However, the term information theory could evolve someday to mean just what many of us would like it to mean. That is the natural way language and science evolves. Information theory, close as it is to mathematical statistics, seems to represent to some people a kind of ideal state for the science of information because they think a field is a science to the extent that its fundamental elements are described in mathematical terms. Measurement is a key element. Agreed upon standards of measurement is another. In that sense, a tailor may be more of a scientist than some of those who would profess to be information scientists. However, pseudo mathematics is not needed (2). The key element is the application of the scientific method. A science is an organized body of knowledge wherein the "scientific method" is applied in order to challenge and upgrade existing theories, postulates, and data.

Now I have an extremely difficult task before me tonight because I could spend hours discussing the nature of science information alone. Obviously, science depends on information -- its accumulation, correlation, modification, synthesis, etc. In an editorial I wrote sometime ago in Current Contents, I spoke about this problem. The editorial was entitled "Who Are the Information Scientists?" (3). It might just as easily have been entitled "Who Are the Science Information Specialists?". The theme of my

editorial was simple. Every so-called laboratory scientist today is the potential information scientist of tomorrow. Scientists have always dealt with information. Their lives have been spent in creating and using it. The important difference is that today many scientists no longer have to create information in the laboratory -- for several reasons. Either someone has already created the information they need, or by suitable manipulation of existing files of information, the desired information can be created. Don't misunderstand me. We will still have to create enormous quantities of new information in the laboratories -- especially in what has been described as the wet sciences (4).

In biology, where life's information systems are so dependent on water, we can characterize information as wet, whereas the information of the physical sciences is dry. Exobiology epitomizes the transition from wet earth biology to dry space biology. In the era of the computer generation of dry scientific information, we are merely making a transition from the old wet laboratory information era. Ironically, the computer that may accomplish this task may well be a wet computer instead of a dry one. Water soluble chemicals may provide the memory units rather than transistors, electronic tubes, and printed circuits. In fact, a whole group of new fields is ultimately concerned with the development of this wet computer -- for example, bionics, bio-medical engineering, artificial organs, artificial intelligence (automata). In short, our wet computer is probably going to be not unlike a computer each one of us now possesses.

And so we come to yet another one of the information sciences -- the study of the brain.

What then is the difference between a laboratory scientist and a theoretical scientist? By my definition, the transition from laboratory scientist to theoretical scientist might easily be a measure of how much wet information the scientist uses in contrast to dry information.

I don't think it should be necessary for me to tell this audience about the growth of scientific literature. There is also a corresponding, but not necessarily proportional, growth of scientific knowledge. In either event, the amount of information in our libraries and in our brains has increased to the point where fantastic opportunities exist for the theoretical scientist or, if you prefer, the science information specialist.

Is Information Science?

I've told you about people who deal in science-information -- the information of science. What about the second facet of my talk -- Is Information science? While I did not earlier take the trouble to be very precise about a definition of the word science, it is going to be necessary for me to do so in discussing the term information. Information is a strange and wonderful phenomenon. The word information, you might say, is full of information -- like talking about "what is the meaning of meaning?"

Since I have been talking about information science, you might reasonably expect me to be able to be very precise about the meaning of information, especially if I am going to claim that information science is

a legitimate term. If I say it is the science of information, then why is information so vague?

Information is an entity apart from the means by which it is processed, the symbols by which it is represented, or the responses made to it. Information is in some respect like numbers in mathematics. It expresses relationship. One could say that the first binary representation in life was made when God created, in the beginning, heaven and earth. Had a Boolean mathematician written the Bible, he might have opened with: In the beginning God created plus and minus, or zero and one. In so doing, he created a means of conveying information. God then went on to create a number of additional binary information states such as night and day, man and woman, good and evil. Boolean formulations happen to be convenient ones since most digital computers are constructed on the binary principle. But surely an equally good case might be made for a ternary or quarternary basis of information. Detailed consideration of this problem is, indeed, the realm of pure mathematics, and I only bring it up here because I don't think we can artificially separate information theory from information science for an indefinite period. It just happens to be more convenient to do that for the moment. Professor Joshua Lederberg has correctly stated that molecular biology, of which I shall say more later, is a corollary of information theory (5). Similarly, information science is also a corollary of information theory.

In an extremely interesting discussion of information, Paul Henshaw (6) last year discussed "Information Per Se" but never grappled with the question "What is information?" He did state that information can be regarded as existing in pieces or concepts but probably not as units. He also said:

> "Information can be transmitted and stored by means of impulses and symbols, but it exerts an influence only after removal from the encoded message and interaction with other information, as can happen in a cell, a brain or a computer.
>
> "Information, as such, is not amenable to quantitation in a manner comparable with that which applies to entities having force, mass or color.
>
> "Information consists of data, especially unrelated facts or statistics."

I am sorry to disappoint those of you who may have been led to believe that I would, in fact, give you a better account of what information is. It is easier to say what it is not. But whatever it is, a very respectable number of respectable scientists, and I include myself in that category, feel that information is a phenomenon common to all the classical categories of science; and, therefore, it will be of some value for you as information retrievers and disseminators and organizers to have brought together in one list some major areas of classical scientific interest that share this highly accentuated interest in this unknown quantity called information.

The remainder of my lecture might well be entitled "From Euphenics to Euphoria." An alternative title might be the "Grammar of Life and Science."

My task is not easy because I must not only be brief and

succinct, but I also must assume you know enough about the various classical fields, to which I will allude, to see the thread of my main argument.

Many of you have heard the story about the astronaut who landed on Mars and encountered a building that looked very much like a computer installation. At the computer console sat a very attractive information specialist. To one side, a group of people were waiting at the end of a long chute. Our future librarian, the information specialist, was busily pressing buttons. In a few moments, the cry of a beautiful perfectly formed 15-pound baby was heard, and the crowd cheered. The rest of the story involves a somewhat humorous comparison of the Martian and Earthly methods of procreation. The main point, however, is that the Martians also used the same computer to retrieve ideal documents.

The ideal baby had all those eugenic characteristics desired by society. Eugenics is concerned with the selective breeding of the genotype. However, the baby also had the ideal euphenic specifications stipulated by his "parents"--the Martian community. Euphenics, a term invented by Professor Lederberg, concerns the selective breeding of phenotypes--of individuals (7). Euphenics is concerned with the complex grammar of life--the genetic black box which generates phenotypes.

In a retrieval system, the ideal document will presumably be the document that contains precisely the information that the requestor wants. For all intents and purposes, the ideal computer is the machine H. G. Wells had in mind when he wrote about the World Brain. In a recent paper, I discussed why the citation index is the first giant step on the way towards the World Brain (8).

The ideal computer system, presumably, will be able to supply ideal documents, that is, documents containing precise answers to ideal and precise questions. Presumably existing computers could do this now. It's the non-ideal question (and to a lesser extent document) that plagues the information systems designer.

Molecular biology, about which you have heard so much, is the fantastic inbreeding of biochemistry with classical genetics and the physical sciences. Some of you will even find the term DNA more meaningful than molecular biology because the term DNA or the term desoxyribonucleic acid has been around much longer than the term molecular biology. And surely, you have heard that molecular biologists are trying to unravel the so-called secret code of life. This phrase "code of life" has a double meaning because it refers simultaneously to our future ability to conceive ideal babies as well as to our future ability to alter or control life as expressed in phenotypes like you and me. In short, molecular biology will affect our understanding of all of "life's" processes including birth, growth, disease, aging, and death.

Now as the third ideal information state, in addition to ideal babies and ideal documents, I should like to refer to that human function called language, the study of which is called linguistics. Linguistics is concerned, in part, with the writing of grammars. Grammars are rule books which tell us how to write ideal sentences. Given a sufficiently detailed grammar, one could theoretically write any well-formed or grammatical sentence. The ideal sentence, however, would not only be well-formed but also would be meaningful and unambiguous. The relationship of linguistics to information science has become fairly obvious so I don't intend to elaborate this point in great detail.

Modern linguistic theory is concerned with the fantastic ability of man to use his brain to command his tongue to produce the right utterances under a given set of circumstances. These fundamental linguistic processes will be better understood as we gain greater insight in molecular neurobiology.

A fourth ideal state necessary to my argument is the domain of chemistry--the chemical compound. The study of chemistry, I submit, is the study of the grammar of chemistry--the rule book, the cookbook, the grammar of chemical synthesis. Obviously molecular biologists are very much interested in the grammar of biochemistry, how protein is synthesized, and they are aware that it is a fantastically complex grammar at that. But they should not be discouraged because it is naive to assume that we know more about the grammar of English than we do about the grammar of biology simply because we have been studying grammar for centuries and only recently began the study of molecular biology. Indeed, in the few years that I have thought about the relationship of genetics to linguistics, I have become convinced that it will probably be the molecular biologists who will revolutionize linguistics and not vice versa. I used to think otherwise.

As a brief aside, I might mention that Professor Saul Gorn has used some interesting terminology to describe components in information systems. He calls people, machines, and programmed procedures--information processors. Biologists would add to this category all lower animal organisms. Professor Gorn would also say that these information processors operate under the control of so-called command languages.

A fifth ideal information state, and the last I will discuss tonight, concerns an area of which I have said nothing so far. It is the ideal psychological state. According to one's preferred terminology, it is, on the one hand, a problem in psychology and, on the other hand, a problem in psychoanalysis. The greatest problem man faces in the near future is the problem of choosing. Science has forced us into the uncomfortable dilemma of having to make decisions of choice. Picking and choosing an ideal baby involves some very serious soul searching. Picking and choosing an ideal document involves some very serious mind and library searching. This will be increasingly truer as our libraries increase their ability to provide detailed information rapidly and cheaply--as in fact libraries become information centers. But in man's quest for knowledge, the most important information is individual self-knowledge. Given the prospect of indefinite prolongation of life, which is a topic of serious consideration today, communal and self-knowledge become vital. What, after all, is it that one wishes to preserve? Deep and careful probing of what we are, or what we want to be, is vital. Since the ideal state of self-awareness is the presumed goal of psycho-analysis, it is not surprising that information processing is an activity of increasing interest to neurobiologists, psychologists, and psychoanalysts--or to use a more general term, the behavioral scientists. In this context, I would refer you to the work of James G. Miller (9) on information overload. In short, much of the mental disease we cope with today is a result of individual and social inability to cope with information.

In conclusion, I will quote from Henshaw--not because I didn't say essentially the same things many years ago, but because quotation does

have a way of giving one's statements a bit of authority, as Professor Steinbach said recently in Science (10). It also gives you the courage to say things for which people may call you a fool. In short, few of us like to be alone in his beliefs.

> "Information science is bursting on us, andwhile information science is not being heralded by events as spectacular as the detonation of bombs or the orbitting of satellites, its influence, it appears likely, will have an impact fully as great, and perhaps much more so.
>
> "The theme of information science remains the same. It is how systems, living or non-living, regulate themselves, reproduce themselves, learn and evolve." (11).

If this sounds somewhat cybernetical, it is no accident. In this connection, it is important to remind you of Norbert Weiner's concept of telegraphing a man (12). In this mechanistic view of man, a completely precise informational profile of an individual could be encoded and transmitted across a wire or other communication channel.

Before I step down, however, I should like to say a few words in a somewhat sentimental vein. My long association with librarians has always made me feel comfortable in their presence. I frequently feel uncomfortable in the presence of mathematicians because we speak different languages. On the other hand, my familiarity with librarians has bred a loving contempt for their complacency and smugness. **On the one hand,** I have the image of the public librarian fighting for freedom of the press; on the other hand, there is the image of the frustrated individual whose only concern is keeping books on shelves and collecting overdue fines.

As librarians, many of you have a right to resent stereotyped images, but you also have to admit that this image is not the creation of the novelists. Unlike scientists, whose public image has been largely a creation of the science fiction writers, librarians have themselves to blame for much that has happened or not happened to them. For example, a new director of an important library was recently appointed. I heard no howls from librarians that this task required a professional librarian of stature and experience; or for that matter at least a man who had devoted himself to one or more aspects of the information problem in a unique way--in short, an information scientist.

During the past decade, I have observed librarians quivering and quaking at the thought that machines were going to replace them. Instead of responding to the challenges of automation, many librarians were too busy defending the notion that librarians and library science had long ago considered the fundamental problems of classification--that there was really nothing new under the sun, whether it be called documentation (which used to be a very dirty word in this building) or information science.

Who among you has the courage to really re-think the library problem--to recognize that information science is a much bigger problem than the document handling business known to traditional library science--in short, which of you is resilient enough to change your name--to call yourself, if you will, an information scientist. Try it out for size the next time you are asked what you are. You may find that in the

explanation you will have to give, you will have made a new place for yourself and your profession--both in your own mind and in those of your fellow human beings. Sooner or later someone has to take a stand. Surgeons were once called barbers. Obstetricians were once called midwives--in some places, midwives are still popular, but who in this audience would use one if given the choice to use a doctor. A rose is not a rose by any other name--names are important.

If my talk here tonight has convinced only one or two of you that the vistas of information science are indeed exciting and close to man's deepest aspirations, then I will have been well rewarded for my effort.

My closing remarks are directed to those of you who are not specifically concerned with the natural sciences.

In 1952 Gerald D. McDonald, Chief of the American History Division of the New York Public Library wrote:

> "I like to believe that in the field of the humanities the library is, and will continue to be, the dominating scene of advanced research."

He also said:

> "...the more important phases of his (humanist's) own research will take place in the laboratory or in the field, with frequent trips, perhaps, to the United States Post Office." (13).

I am not in a position to comment as to whether, during the past decade, there has been a greater tendency for humanistic scholars to use the libraries less. I do know many fields of the so-called humanities are so science-oriented that it is almost a rationalization

to persist in the arbitrary distinction of humanistic vs. scientific. I also know that many activities in the humanities involve enormous field projects. I find it rather ironical that it should be the natural sciences, formerly so laboratory-oriented, which today are increasing their dependence upon libraries and sources of information. However, this will only continue to the extent that libraries and information centers remain fruitful sources of information. It was, among other reasons, the failure of libraries to adjust to the changing needs of the humanities and other research areas that literally drove people to seek solutions outside the library. Fortunately, that situation has drastically changed. We should all be proud of the pioneering roles that Drexel Institute and the University of Pennsylvania are playing in the new and exciting domain of Information Science.

REFERENCES

1. L. Wilson, "Information Scientist Defined," (Letter to the Editor), Special Libraries 54(3), 175 (1963).

2. M. Taube, "A Note on the Pseudo-Mathematics of Relevance," American Documentation 16(2), 69-72 (1965).

3. E. Garfield, "The Informatorium No. 2: Who Are the Information Scientists?" Current Contents of Chemical, Pharmaco-Medical and Life Sciences 6(24), 5 (1963); and Current Contents of Space, Electronic and Physical Sciences 3(26), 4 (1963).

4. P. Henshaw,"Information Per Se," Nature 199(4898), 1050-1052 (1963). See esp. p. 1051.

5. J. Lederberg, "Molecular Biology, Eugenics and Euphenics," Nature 198(4879), 428-429 (1963).

6. P. Henshaw, op. cit.

7. J. Lederberg, op. cit.

8. E. Garfield, "Science Citation Index--A New Dimension in Indexing," Science 144(3619), 649 - 654 (1964).

9. J. G. Miller, "Sensory Overloading," Psychophysiological Aspects of Space Flight, B. E. Flaherty, ed., (New York: Columbia University Press, 1961), pp. 215-224. See also J. G. Miller, "Information Input Overload and Psychopathology," The American Journal of Psychiatry 116(8), 695-704 (1960).

10. H. G. Steinbach, "The Quest for Certainty: Science Citation Index," (Review), Science 145(3629), 142-143 (1964).

11. P. Henshaw, op. cit., p. 1052.

12. N. Wiener, God and Golem, Inc., (Cambridge: M.I.T. Press, 1964), p. 36.

13. G. D. McDonald, "Application and Limitations of Subject Headings: Humanities," The Subject Analysis of Library Materials, M. F. Tauber, ed., (New York: Columbia University School of Library Service, 1953), p.55.