Transcript of talk at Industrial Dinner. Edited later by D.R. Swanson for publication by ICA

"Communication of Scientific Information"

Dr. Don R. Swanson, Dean
Graduate Library School
University of Chicago

Mr. Wright, ladies and gentlemen, your very generous introduction almost makes me feel as though I were here on purpose. Mr. Licklider, with whom I am well acquainted, is brilliant, imaginative, and entertaining. It is indeed regrettable that he was unable to be here this evening, and I think you should not be fooled by Mr. Wright's attempt to make this calamity seem almost like a stroke of fortune. I am afraid it reminds me a little bit of President Nixon's awesome attempt to characterize the war in Viet Nam as possibly being our finest hour. My disappointment is keen -- not that I have to speak tonight, because it's an honor and a privilege. But look at it this way -- I have to come here to listen to me instead of to Mr. Licklider. To make matters worse, and I hope you can appreciate my position with regard to some students at the University of Chicago, because many weeks ago I assured them that it was indeed well worth $12.50 to come here this evening. So yesterday, as soon as I knew, ahead of everyone else of course, who the substitute was going to be, I hastened to advise them to try to sell their tickets very quickly while there was still a chance. In fact, it turned out okay for some enterprising students in the Business School. They even managed to sell short, before the news hit the market. Mr. Wright only partly conveyed to you the sense of crisis I'm sure he felt yesterday. You see, in casting his net far and wide for a substitute speaker, if you can envision casting a net when you're trying to scrape the bottom of a barrel, he enlisted the help of others who made a number of frantic phone calls -- one of which reached me. I then called Mr. Wright back, whom I had never met before, indicating that I would be pleased to speak at this banquet. He replied, in the following order, 'Wonderful! I'm delighted! Who is this?'

Now what can I tell you about libraries, nodes, networks, the knowledge explosion, computers, and so on. Picture if you will, decades hence, a dozen trillion-bit millisecond access memories, backed up by billion-bit nanosecond storage, hooked into multiple-access computers with remote consoles, all tied together with a far-flung electronic communication network, the whole constituting a versatile multi-user interactive on-line rapid response system. We all know this is the answer, and it might be fun to dream of what we could do with such a system. There is a cartoon, showing Charlie Brown carrying a sign, "Christ is the answer." In the last frame, Snoopy carries a sign, "What is the question?" This is my topic.

I take it for granted that we are in some kind of trouble; there is too much information. The ever-increasing record of the present that Mr. Tuthill referred to may be the problem, though if we look at the percentage increase per year and realize that so are many other things increasing at a certain percent per year, including our resources for coping with the knowledge explosion, it isn't all together clear that one can make a good case for that being the core of the problem. We are faced, however, with a fundamental limitation, probably, in the rate of input of information to the human brain. When we measure that against the relentlessly accumulating backlog of information, then we see that there is indeed a problem -- the backlog continues to accumulate; it doesn't disappear or dissolve. And we are in a situation where there is too much to read, even to the point where it may well take more than a lifetime to work our way through to the frontiers of knowledge in some new field.

I wouldn't characterize it as an information or knowledge explosion, because it's altogether quiet, but it is a silent, relentless flood in any event; and of polluted water at that. And if we are unable to catch up to the frontiers in a lifetime, then we'll

face a situation where we'll be spending the next century rediscovering the knowledge that we've already won this century. Now our concern with the problem is not limited to those of us in the library field or the computer field. There has been a great deal of concern at the national level, going back farther than I think most of us are aware. The legislative acts, which established the Office of Naval Research in 1946, the Atomic Energy Commission in 1947, the National Science Foundation in 1950, and NASA in 1957 all contain provisions for what these organizations should do to make sure that the scientific and technical record which they generated was efficiently used. All of these organizations now have large information programs as well as many other government and federal agencies including particularly the three national libraries -- the Library of Congress, the National Library of Medicine and the National Agricultural Library.

Interest in the part of Congress and the Executive Department continued unabated between 1955 and 1960. Humphrey's Subcommittee on Reorganization and International Organizations of the Senate Committee on Government Operations held many hearings, published a great quantity of excellent information that brought together commentary of experts on the problems of scientific and technical information.

The Committee on Science and Astronautics also held hearings and took considerable interest in the problem and issued reports and legislative proposals. Beginning in 1958, the Executive Branch stimulated a number of plans. A panel of the President's Science Advisory Committee produced a report under the supervision of W. O. Baker. In 1962, another panel of the Science Advisory Committee under Weinberg reported on this problem. In 1962 also, a task group under the Science Advisory Committee under J. H. Crawford, as a result of the Weinberg report, the

Committee on Scientific and Technical Information, COSATI, of the Federal Council of Science and Technology was established to coordinate the various federal programs of information. In 1964, the Office of Science and Technology set up an adhoc task group under Licklider that produced a study, and most recently, under the auspices of the National Academy of Sciences and the National Academy of Engineering, the Committee on Scientific and Technical Information called SATCOM was created, and has recently produced a rather comprehensive report. Mr. Licklider and I were two of the 15 or 20 members of that committee. He and I were also the only ones on the particular subcommittee. We never managed to meet as that subcommittee, so I wrote my section of our report and he wrote his section, and our chairman, undaunted, published both sections in different parts of the SATCOM report.

Now I'm going to direct some of what I say now to the spirit of the Baker, the Weinberg, and the SATCOM report as I perceive biased and distorted. That is, the spirit of these will be partly reflected in what I say, but I make no claim that it necessarily reflects the intent of their authors. I have read these reports, been stimulated by them, been influenced by them, and I come away with a number of dominant impressions on where the nation's thinking is headed with regard to the problems of scientific and technical information.

One of these is that we have a pluralistic system and it's here to stay, and it should stay. We have thousands of institutions and services, libraries, indexing and abstracting organizations that contribute to our attempts to control and provide access to scientific and technical information. This system, such as it is, has evolved spasmodically through adaptation to locally perceived needs and opportunities. In short, it's a cluge, but it's ours and it has advantages over what is often proposed as what I would call the monolithic system -- a big national center somewhere to store, handle and disseminate all scientific and technical information.

Our pluralistic system, whatever its shortcomings, is flexible, adaptable, and supports innovation, initiative, and responsiveness. It offers better opportunities for being strengthened, and for taking positive action to cause it to evolve in the right direction than does the idea of scrapping everything and starting over with a big new mechanized center and the trillion bit memories that store everything. But we have to have a good concept of what this all is before we can see the kind of action to be taken. How can a particular library or information service transcend its local limitations? The concept that is central to this idea is, of course, that of a network. The image of our information system should be the Bell Telephone System, not Fort Knox. The problem can be regarded in two major portions: the first dealing with bibliographic control, intellectual access to information, or catalogues, published indexes, and so on. How can these function as a system? When we think of the catalogue of an individual library we have something that, in Mr. Licklider's terminology, is a node, in this case, without a network. But the concept of the Union Catalogue, that is, a catalogue that lists the combined resources of many libraries, is central to the idea of network. Of course, the indexing and abstracting services transcend the boundaries of local institutions, and cover material that is widely held and widely dispersed.

These are obvious facts. I'm sure you've all heard and are fully aware of both of these. I invite your attention to them at this point because of the way they fit into the notion of a network, and because it is only by recognizing these key elements that we can eventually find the points of leverage for expanding our present cluge into a true network.

The other half of the problem, besides the bibliographic access, is of course the text itself. It is less easy to see what the significance of this is insofar as the network concept is concerned. Perhaps, historically, we could look at the idea of

interlibrary loan as an attempt by individual libraries to provide resources and access that they can't by themselves, through the sharing of resources with other libraries. But in principle, in any event, the idea that no one institution can afford to be altogether self-sufficient suggests the need here for some kind of sharing and some kind of communication and transportation of materials. Here the notion of a central store might be appealing; the stores' trillions of bits and provides electronic access remotely. But let me point out, too, that there is a competitor to this idea. If only we could make cheaper copies of everything and distribute this storage all around the country, it might cut communication loss, so we have a competition between whether to store things in one place or make a lot of cheap copies and store them in many places.

Now if as a computer engineer I were to specify that one of the things we need for our system is a two hundred million bit storage device that fits in about a cubic foot, is portable, requires no equipment in order to gain access to it, can be accessed in a matter of minutes and costs less than $10 or $20 and is producible in thousands of copies, distributed to remote points where users can have ready access. You see, I have essentially reinvented the book, and not altogether facetiously.

I attended a conference a few weeks ago in San Francisco of the American Society for Information Science. They had many exhibits at this conference of on-line interrogation, via teletype, of remote computer-stored information. This was very interesting. The dialogue that one carries on with a store was fun to play with. It is of great significance so far as future information systems and future networks is concerned.

Yet, one should not overlook the engineering comparison between interrogating a computerized store, by way of teletype, with flipping the pages of a book in order to retrieve the same information in printed form. This isn't to say that the book can't

always compete because there are certain logical combinations of terms and fancy things one can do in searching with a computer that one cannot do with a book. Yet, a great deal of what was done with those stores, and a great deal of what needs to be done can also be done with this competing idea. And the competition is not a matter of principle in any sense; it's a matter of systems planning and purely engineering economics. So superimpose now on these two basic ideas of intellectual access on the one hand and the sharing of textual resources on the other, either through computerized stores or through better dissemination of cheap storage devices, such as books. We need to superimpose good communication networks, to be sure; and secondly, coordination and cooperation. That's a very difficult notion to deal with, because society just isn't organized properly to cope with our information services. Everything cuts across the boundaries of everything else, and we have to depend upon voluntary coordination and cooperation rather than upon establishing new organizations and seeking to crystalize lines of responsibility and authority. It is difficult to work informally, and by means of coordination and cooperation; yet that seems to be the only alternative at our disposal, in view of the greater difficulties that it seems to most of us would be encountered should we try to design the single centralized monolithic system.

Apart from recognizing the conceptual nature of the problem, let's notice that many things of great importance that are presently going on fit well into this framework. I've already mentioned the three national libraries. Let me further mention what many of you may know about the programs in which they are engaged, addressed specifically to the network concept. The John Crerar Library will become a Regional Medical Library Center. Under the sponsorship of the National Library of Medicine,

one eventually will be able to carry out machine searches. The MEDLARS program of the National Library of Medicine is a national effort. There are a number of such regional centers, the products that they produce are widely distributed, and this is a means of coordination and the provision of services of a national character that are accurately described in the framework of a network; similarly, the programs of the Library of Congress, the National Agricultural Library, of NASA, the AEC, and a variety of others including many local public library networks.

There is a lot going on. It's beyond the scope of what I wish to say here to review the field for you, but only to notice that having presented the notion, there is much evidence of progress in spite of the impression that we have as users of these systems that they tend to function independently and rather chaotically. There is, nevertheless, good evidence of progress toward network functioning.

Now I mentioned a few minutes ago that the rising flood of information consisted of polluted water. What can we do about the quality problem? This is central. I think we all have the impression that much of what is written should never have been written and that we could undoubtedly throw away at least half of what's in our libraries now, if only we knew which half. In fact, the question probably isn't whether we should throw away half or not, but rather whether we should do it more rationally than the way we're now going about it; namely, by making it relatively inaccessable by being cluttered up and insufficiently systematized.

This question of quality control emerges again and again. I can remember discussions on various committees and panels ten years ago. Bill Baker agreed that this was central at the time the Baker report was written, and there is much concern

about it. But in the spirit of this philosophy that we should look to our present pluralistic system, see how it functions and see where it needs strengthening, let's do that for a moment with regard to the quality problem. Let's ask do we now have any measures for quality control, and can these reasonably be strengthened.

As soon as we ask that question we begin to see that there are indeed such measures. The scientific community does sit in judgement on its own product. How does it do this? First of all, there are referees for scientific and technical papers. The more prestigious journals in particular are more rigorous in the refereeing practices. I will admit this does not keep things from getting published. There is some evidence that if a paper isn't accepted in the prestigious journal it goes to the next choice of the author and so on, until eventually it gets into print. But something else happens to it, apart from the refereeing system, if it doesn't get into one of the central journals; namely, it gets cited less. An author puts a list of references at the end of a paper, and his paper in turn is cited by others. And while the notion that the number of times a paper gets cited has evoked expressions of dismay and shock that one should regard that as somehow related to the quality of a paper, yet all of the sociological studies of citation counts that have been done indeed support this notion -- that the good papers are the ones that are frequently cited.

Papers published by prominent scientists of high quality are more accessable than papers of lower quality because they are more often cited, and because this notion of following chains of citations and references is such a common practice. It's the dominant way that scientists, in fact, gain access to information.

In addition to citations there are, of course, reviews written. The notion that in order not to perish one must publish has a counterpart --- that it is possible to perish by publishing low quality stuff, because it gets exposed to the critical gaze of one's colleague's, and reviews, as some of you may have experienced or known, can be scathing. The fact that something might possibly be reviewed, I think, has a salutary influence on the quality of what is now published. As bad as we might think it is in general, this influence is at work. And the argument for recognizing this influence is, of course, it helps us to think of ways to stimulate and encourage such influences.

Not only are critiques written, but reviews and summaries tend to counteract the information explosion. A good review, hopefully, will make one article grow where ten grew before. Unfortunately, most reviews are really not of that character. But the deficiency, I think is being more widely recognized. Reviews now tend to be more bibliographic guides to the literature, and it's a rare one that really replaces what went before, though this does happen and I think particularly in the monograph literature; but it's a beneficial trend that deserves acceleration.

Well, the foregoing influences partly work, and partly they don't. I have no specific solutions to propose. I think the best step that we can take at this point is to make explicit what the good features of our system are in order to give us a point of attack, and hopefully enable us to strengthen it. Our whole system now is enormously complex and gives every indication of becoming more so. And this complexity is pertinent to questions of cost and to economics. Our methods of financing information services are simplistic and ill-conceived. There is little

relationship between cost and value, and even less between who pays for and who receives the services. Now a conspicuous exception to these particular remarks is, of course, The John Crerar Library; but I wish to point out that it is an exception, and these remarks are specifically addressed to the hope that the kinds of services that it represents in its mode of operation may be a key element in strengthening our entire national network. How often would we, for example, welcome the opportunity to pay the necessary price for extra services, because the only thing that really stands in the way of achieving this imagery that I painted of the trillion bit stores and the rapid access and so on is money. There are no limitations of technology that prevent us from doing anything that we want to do. But how do we bring the money to bear at the right point? It seems to me that what we need is the test of the market place. Then we could begin to understand more about the value of information services. With this understanding, not only can improved information systems be created, but perhaps computer manufacturers might be encouraged to develop equipment better suited to those systems. How does the library field tend to operate now? It rather resembles a Marxist philosophy of to each according to his need. This may well be outmoded and should give way to a more flexible, responsive, and realistic consumer-oriented philosophy, and one in which we might at least begin an elementary analysis of effectiveness as a function of cost.

Now there is a seeming anomoly in the whole present state of affairs, notwithstanding persuasive evidence that existing libraries and information services are inadequate at best; yet they are used in such a way that only a small part of their potential value is realized. The typical users, and again I suspect the users of John Crerar Library are in exception, the typical users' perception of these

services is marvelously simplistic. Many studies of this have been carried out. I am not giving you unsupportable impressions, though in part I am abstracting impressions from these studies, but they have shown that library users do not understand the construction of major location tools, have little concept of the multitude of reference and bibliographic services relevant to their own specialty, and do not even perceive the librarian as someone who might provide assistance in leading them through the maze of potentially pertinent information services.

Well, if this is so, and if information services are not being used, can we not simply conclude that they aren't needed? Perhaps. This is a matter of conjecture. I'm inclined to think that present services aren't being used because they aren't well adapted to need, and this is so because the machinery of adaptation suffers greatly from nonexistence. Things not viable have been allowed to live, and much needed services have never been born. A close relationship between price paid and services received might well provide the missing mechanism of adaptation.

However, information systems are cumbersome and slow to respond to forces of change. What seems inevitable is that they will grow more complex and offer scant comfort to those who are awaiting a simpler life, to those who are awaiting the push-button, remote access, rapid response, multi-access, on-line system. I think it is true now, and it will be true for a long time to come, that anyone who himself wishes to use effectively the world's information resources will have to become, and remain, far more familiar with the apparatus of communication than is the typical scientist or engineer nowadays.

So I've tried essentially to focus on what is the question. But I must admit that the very manner of asking it has reflected a bias and implies that I have certain preferences, at least, in approaching a solution. See, I feel that the proper role for computers, consoles, rapid access, large capacity storage, and electronic communication all will emerge naturally, and of its own accord, if we solve the more basic essentially economic and sociological questions first, so that then the environment will become more hospitable; that is, more profitable for technological change.

Now I sense that I have reached a point in my talk where I see a number of new resemblances to the war in Viet Nam -- it's been too long, it was miscast from the beginning, the whole objective is less clear looking back than it seemed at the time it was started. Obviously, in substance it's over with, but it's very difficult to find a graceful ending. But on the chance that I may have misunderstood my mission, and your interests, after my conversations with Mr. Wright and Mr. Budington, I'll take a clue by drawing on a story about a friend of mine, an incident that took place at his wedding. This friend I shall call ~~Sir Gay~~ Sergei (!), because that's his name. His native language was Russian, though he spoke English fluently. And yet, when the minister said, 'Wilt thou take this woman to be thy lawful wedded wife?' the wilt, the thou, and the thy were quite new to him; and there was a long, embarrassing pause while he sorted these out in his mind. And finally ~~Sir Gay~~ Sergei said to the minister, 'Would you please repeat the question?' So I'll have to put that to Mr. Wright, and if there is more that he would like to know from me the question will have to be repeated.