R63a

Science, Government and Information: 1988 Perspective[1]

Alvin M. Weinberg

Senator Hubert Humphrey was the unsuspecting godfather of the 1963 President's Science Advisory Committee (PSAC) report: "Science, Government, and Information: The Responsibilities of the Technical Community and the Government in the Transfer of Information."* His Senate Subcommittee on Reorganization and International Organization had for several years been promoting an all-encompassing Department of Science. A major justification for such a department was the alleged disarray of the country's scientific information apparatus. Were all government-sponsored research placed in a single department, deficiencies in communication could be dealt with globally instead of piecemeal--at least this was the rationale offered by the late senator and his enthusiastic staff for a Department of Science.

The executive branch, under both President Eisenhower and President Kennedy, would have none of this. Putting all of government science under a single agency made no more sense than putting all accounting, or indeed, any other overhead activity under a single agency. Most government science is aimed at accomplishing a non-scientific mission, like stronger defense or better health. Thus, for example, to split military R&D, a means for helping to achieve better defense, away from the Defense Department simply was wrongheaded. If scientific communication was deficient, fix it--don't start a new department. The President's Science Advisor, Dr. Jerome Wiesner, therefore responded to Sen. Humphrey's challenge by establishing a panel to study scientific communication, both inside and outside the government. I was asked to chair the panel, not because I had strong credentials as an expert in scientific communication, but I suppose, because my credentials (Director of Oak Ridge National Laboratory) were not clearly inferior to those of any other member of PSAC--that is, with the exception of Dr. William O. Baker. He had chaired PSAC's first panel on scientific

---

*I shall refer to the report as SGI.

[1] Presented as the Joseph Leiter lecture, National Library of Medicine, April 7, 1988.

information. This panel had, in 1958, recommended the establishment of the Office of Scientific Information Services in the National Science Foundation. Dr. Baker at the time was soon to become an emeritus member of PSAC, and would then be technically ineligible to chair a panel. However, he was one of our most active and helpful panelists. The remaining panelists included two Nobel Prize winners--Eugene P. Wigner and Joshua Lederberg; a Krupp Energy Prize winner, Karl Cohen, chief scientist of the Nuclear Division of General Electric; John Tukey, a distinguished statistician from Bell Labs; Jim Crawford, a solid state scientist from Oak Ridge who later chaired the Physics Department of the University of North Carolina; Lew Hammett of Columbia, a past-president of the American Chemical Society; Andrew Kalitinsky, a senior aeronautical engineer at General Dynamics; G. W. King, an information scientist at ITEK; W. T. Knox, an information scientist at Esso Research & Engineering Co.; and Milton Lee, executive director of the Federation of American Societies for Experimental Biology. Francois Kertesz, a chemist who was involved in information science at Oak Ridge, served as rapporteur and informal adviser.

Of the twelve panelists, three were physicists, four were chemists, two were mathematicians, two were biomedical scientists, and one was an engineer. Only three (King, Knox, and Lee) were professionally involved with handling of scientific information, though several others (Baker, Tukey, Crawford, and Lederberg) had given considerable thought to the problems. The rest of us were working scientists/engineers, or research administrators: our approaches to the issues were therefore hardly encumbered by much prior knowledge.

Our panel met every month for about two years. At each meeting we were briefed by professional scientific information experts both from government agencies--including the National Library of Medicine--and from professional societies. During the summer of 1962, I took leave from my job at ORNL, and moved to Washington to draft our report. I would write every morning in my office at EOB; I would then lunch with one or another of the many information specialists in the area; and then I would rewrite my morning's draft to incorporate what I had learned at lunch. By the end of the summer, my version of the report was finished. Our panel then met for another half-year arguing over each point in the original draft; the final draft was then reviewed by the entire PSAC, and was issued in January of 1963.

Two points stand out particularly in my memory of the genesis of Science, Government and Information: first, the great impact of Josh Lederberg's felicitous emendations, corrections, improvements, insights; and John Tukey's insistence that we "de-crisify" the report. I had chosen the title "Science, Government and the Information Crisis;" John argued that the issue would be with us forever and therefore "crisis" was an inappropriate word. I am still grateful to John for his measured wisdom in removing the original exaggerated sense of crisis which permeated our first draft.

What Science, Government and Information Said

The 51-page report ranged widely: philosophic disquisition on the nature of the scientific enterprise; characterization of the "information problem"; advice to government agencies and to the technical community; and less specifically, advice to librarians, documentalists, and individual scientists. Altogether the report contained 11 major recommendations. The tone of the report was exhortatory: its recommendations were couched in simple declarative sentences. As I reread it now, I blush at its pontifical tone--but after all, it was a report carrying the imprimatur of the President of the United States!

Two main themes dominated: first, that handling of scientific information was, or ought to be, an integral part of science; and second, that retrieval of information was not the same as retrieval of documents. The first tenet was stated in the opening two paragraphs of Science, Government, and Information:

> "Transfer of information is an inseparable part of research and development. All those concerned with research and development--individual scientists and engineers, industrial and academic research establishments, technical societies, government agencies--must accept responsibility for transfer of information in the same degree and spirit that they have accepted responsibility for research and development.

> The later steps of the information transfer process, such as retrieval, are strongly affected by the attitudes and practices of the originators of scientific information. The working scientist must therefore share many of the burdens that have been traditionally carried by the professional documentalist. The technical community must devote a larger share than heretofore of its time and resources to the discriminating management of the ever-increasing technical record. Doing less will lead to fragmented and ineffective science and technology."

This statement, especially the last sentence, bespeaks a certain hubris: who can say whether science and technology are too fragmented and ineffective. Yet the specific recommendations--e.g., to make the handling of information a part of an agency's R&D structure, not of the administrative structure, seems to me still to make good sense. Nor can I fault such exhortations to scientists as appeared later in the report as "Write more clearly"; "Write better abstracts and titles"; and "Spend more time writing thoughtful review articles!"

The second theme--that retrieval of documents was not the same as retrieval of information owes its prominence in the report to briefings by Prof. Touloukian, head of the Thermophysical Properties Center at Purdue University. With the technology available in 1962, our panel could not conceive of super computers sensitively compacting the literature with little or no human intervention. We therefore hit on the information center, manned by scientists as well as information specialists, as central to rationalization of the information system. Indeed, SGI's main legacy is still regarded as its visualization of the information center as a key, if not the key, to a future information system. The panel's position on information centers was stated as follows: "The Panel sees the specialized information center as a major key in the rationalization of our information system. We believe the specialized information center should be primarily a technical institute rather than a technical library. It must be led by professional working scientists and engineers who maintain the closest contact with their technical professions and can make new syntheses with the data at their fingertips."

We came to this position partly because of our puzzlement over the difference between documents and information: information centers, as contrasted with libraries, purveyed information rather than documents; and partly because the chairman of the panel had first hand experience with information centers at the Oak Ridge National Laboratory. During the 1960s, the laboratory housed several information centers--perhaps most notably Kay Way's nuclear data center. This center provided annotated and standardized nuclear energy level schemes and cross-sections for all known nuclides. K. Way herself had started collecting the data at the Chicago war-time Atomic Energy ("Metallurgical") Laboratory, and had continued the work at the National Research Council before moving to Oak Ridge. Other centers at Oak Ridge dealt with radiation shielding, and reactor safety. In all cases the staffs of the centers were professional scientists who spent at least half of their time digesting, evaluating, and summarizing the mountains of data that flowed from the many laboratories working in these fields. And the report repeatedly stressed that the Information Center was part of a research laboratory, not part of a research library.

The report leaned over backward in its assessment of automation: "Mechanization can become important, but not all-important." Remember this was before on-line terminals were common, and personal computers were not even a gleam in a very young Steve Jobs' eye! Though we sensed that automation was coming, perhaps we downplayed its role because we were so anxious to elevate the position of the scientist in the information transfer chain.

Other recommendations dotted the pages of SGI: Citation Indexing was a favorite, largely reflecting an enthusiasm of Josh Lederberg and John Tukey; resident referees for government laboratories were proposed as a means of imposing higher standards on government contract reports; and we suggested that government departments that dominated, but did not monopolize, certain fields (as, e.g., NIH in biomedicine, or AEC in high energy physics), be designated "delegated agents" for controlling and organizing the flow of information in these fields. This suggestion had a bit of the flavor of Sen. Humphrey's Department of Science: information in each broad field of science would be the responsibility of a delegated agency, even though research was performed in the field by several agencies.

Though these, and several other suggestions, abound in SGI, the main focus of the report was its insistence that scientific information really was part of science; and that the Information Center, staffed by working scientists, was a palliative if not a panacea for the information "crisis."

How SGI was Received in 1963

Though SGI was aimed primarily at the research community and the agencies that supported research, the report had rather little impact on working scientists and engineers. It was not that the scientific community necessarily objected to or even lacked sympathy with the report's findings and recommendations; it was rather that the average scientist seemed to be unaware of the report's existence.

On the other hand, SGI created if not a sensation, then consternation among professional librarians. After all, the word library or librarian occurs not more than half-a-dozen times in the entire report, and generally in a vaguely scolding tone. Librarians were pictured as being somewhat anachronistic, as not keeping up with modern developments in automation or in new methods of bibliographic control. Perhaps most threatening to librarians was the call to the scientific community itself to take over functions that librarians had traditionally regarded as theirs. At least this is the impression one gets from the 1963 meeting of the Special Libraries Association. A committee had been appointed to recommend actions SLA might take in response to SGI. Though several on the committee conceded that SGI had pointed to legitimate deficiencies in the way librarians dealt with the information onslaught, most of the committee regarded SGI's seeming dismissal of the role of librarians as reflecting an unjustified scientistic bias on the part of SGI, and I was told at the time that one indignant librarian ceremonially burned a copy.

I illustrate the flavor of some librarians' reaction to the report with a quotation from Joan Mavis's comments in the July/August 1963 issue of Special Libraries (p. 331), the journal of the Special Libraries Association. "According to the Weinberg Report, the present day librarian or information specialist will have to give up his work to a scientist who will evaluate the work of other scientists while the special librarian steps down to work of a clerical order.

"How such a report can be accepted so easily remains a matter of surprise. It either wholly disregards, or regards with insufficient attention, certain facts obvious to anyone who is working in the field of science, to wit, scientists capable in an area specialty have neither the time nor the interest to be information specialists for the nation.

"Area specialization among eminent scientists is now so deep that 'hard' scientists have become myopic in outlook--the truly impartial evaluators are the special librarians. Not only do they evaluate a scientist's offering on the basis of a wide acquaintance with the other offerings in the field, but they evaluate constantly on the basis of relevancy to the needs of their particular clientele."

Perhaps I ought, 25 years after SGI, to explain how this seeming anti-librarianism crept into SGI. When the President's Science Adviser announced to PSAC that a panel on scientific information was to be established, the late Prof. I. I. Rabi, a distinguished member of PSAC, could hardly conceal his disinterest. "Humph--Librarian's work" was his only comment. In the face of such indifference, our panel had to convince the scientific community that information was part of science--not simply "Librarian's Work." That there were legitimate deficiencies in librarianship 25 years ago could not be denied; but the report was not really aimed at the library community--it was aimed much more at the scientific community and at the agencies that supported science.

The appearance of SGI seemed also to exacerbate the rivalry between the community of documentalists and the community of librarians. On the whole the report seemed to be more approving of documentalists, less approving of librarians--and I sensed at the time that the documentalists enjoyed what they interpreted as approval of their approach to information. Indeed the information center (described in SGI as a research institute rather than a library) seemed close to the kind of organization documentalists considered to be their natural habitat.

The report became quite popular in the academic information community. After all, about a quarter of SGI was devoted to philosophic analyses of the role of information in the scientific enterprise and some genuine issues in the philosophy of science were raised. For example, the report insisted that to understand the handling of scientific information one must understand the very structure of science itself--thus some attention was devoted to how science divides into disciplines, how these disciplines relate to each other, and how they relate to the missions of the agencies that support them. The mission-discipline duality was a recurring theme in SGI since findings in mission-oriented science were often of interest to discipline-oriented science, yet, because the two communities often interacted weakly, transfer of information between them could be laborious and inefficient. I have suspected that SGI's seeming popularity among academics was attributable to its raising such issues, which appeal to professors and students of information science.

The government agencies took SGI seriously; after all, it was a Presidential document. Though no specific action was taken by the Federal Council of Science & Technology, SGI was taken as sanctioning information centers--and in the years immediately following SGI--new information centers sprung into being.

Science, Government, and Information in 1988

During the past 25 years I have used information systems as an administrator, a government official, and a working researcher; but I have not been very close to the professional and governmental information-handling organizations that might have been influenced by SGI. Occasionally, as in preparing for this lecture, I have turned to my friends in the information community to brief me on the current situation--but by and large my impressions remain anecdotal and personal.

I have two main impressions: first, that information centers as conceived in SGI, have not emerged as dominant elements of the information system; and that SGI greatly underestimated the influence of the computer on today's information systems.

As for the specialized information center, some 400 were in existence in the U.S. at the time of SGI. But by 1980 this number had shrunk to only

100 when Maskewitz and Carroll reviewed specialized information centers in Volume 15 of the Annual Review of Information Science & Technology (the existence of such a review journal itself bespeaks the enormous development that has occurred in these years). According to Maskewitz and Carroll, SGI gave initial impetus toward information analysis centers, but they have not generally played the key role as synthesizers envisaged in SGI. Perhaps it was too much to try to impose on the scientific and technical community a new kind of social organization that did not spring entirely naturally, and in evolutionary fashion, out of the perceived needs of practitioners themselves.

Though scientist-dominated specialized information analysis centers had not emerged as the key element in the information system, on-line database services have expanded enormously. This growth is illustrated in Table 1, which appears in the January 1988 issue of Directory of On-Line Databases.

Table 1

Directory Issue	Number of Databases*	Number of Database Producers	Number of Online Services	Number of Gateways
1979/80	400	221	59	
1980/81	600	340	93	
1981/82	965	512	170	
1982/83	1350	718	213	
1983/84	1878	927	272	
1984/85	2453	1189	362	
1986	2901	1379	454	35
1987	3369	1568	528	44
1988	3699	1685	555	59

*Includes distinctly named files within database families.

The databases in the table are not all scientific--indeed, a majority of these are legal and financial databases. Nevertheless, the tenfold increase in on-line services in less than 10 years is astonishing. The

Directory attributes this growth to several technologies and information products, all of which have matured during the 25 years since SGI was published. These include databases, that is collections of numeric or textual material processed for electronic publishing; timesharing computers that permit simultaneous access by many users; interactive computer programs; rapid access storage devices; cheap computer terminals and microcomputers; and telecommunication networks.

To a degree, then, these on-line databases provide some of the services envisaged in SGI for information analysis centers. Generally, the databases do not provide analytical reviews, though in some cases they provide actual numeric data--for example, if I subscribe to the National Standard Reference Data System, I can call up on my desk computer the melting point and specific heat of uranium. Obviously, a great deal of analysis goes into the preparation of such databases. In some cases the databases themselves are inputted by the working scientists. For example, the newly formed superconductivity information system, being developed by the Department of Energy's Office of Scientific and Technical Information, now has about 200 subscribers all of whom are active researchers in superconductivity. They provide abstracts, data, and short progress reports on their research to the central file. In return they have full access to similar information from all the other subscribers. In a sense, then, the superconductivity data system provides relevant information to a dispersed company of researchers who in turn provide reviewed information to the system: much of the function of an analysis center is therefore provided, though in a decentralized manner.

Another, and largely unanticipated development since SGI, has been the rise of serious scientific journalism. The scientific journalist, who is usually a highly qualified scientist turned journalist, now plays an important role as a reviewer and compacter of the scientific literature. Nature and Science today devote much more of their space to thoughtful reviews by professional scientific journalists and scientists than they did 25 years ago. Or, on a more professional level, review journals are now much more common than they were 25 years ago--for example, Annual Reviews, which began originally to review the literature in biochemistry, now reviews 27 separate fields of science and technology. And semi-popular scientific newspapers and journals now are common--whether part of the New York Times, as journals like Scientific American or Discovery, or as a newspaper for scientists, like Eugene Garfield's The Scientist.

Most of these secondary sources have been established by entrepreneurs who recognized an empty niche, and proceeded to occupy it. This is true of the database services, as well as the journalistic activities--these services more and more have had to meet the test of the marketplace.

Conclusion and Outlooks

As I contemplate the vast continent of scientific information systems, I realize that these systems, and their customers, are a diverse crowd. Schemes appropriate for some kinds of users are hardly relevant for other users.

In my view, the users fall, broadly, into two classes: first, those for whom time is of the essence, who operate necessarily in a state of perpetual "crisis" and who seem ready to pay money for their information services; and those for whom deliberateness and studied response is of the essence, for whom haste makes waste, and who, in the customary scholarly tradition, expect information tools to be free services.

In the first category I would put most engineering and operational scientists. Thus, a medical doctor who must diagnose a patient on the basis of various tests, finds on-line medical diagnosis systems extremely helpful. For him time is of the essence; and on-line databases, particularly those embellished with modern artificial intelligence methods are obviously helpful.

Another example comes from my own field, nuclear energy. The accident at Three Mile Island was a prime example of both information deficiency and information overload. The deficiency lay in the failure of the operators at TMI-2 to know that accidents almost identical to TMI-2, though less serious, had already happened at Davis-Besse and Rancho Seco; had the TMI operators known of these, they surely would have diagnosed their problem before the core melted down. Information channels were overloaded: once the accident started, the control room was deluged with a bewildering avalanche of lights, bells, announcements, data. What was desperately needed was a means of analyzing the raw data, extracting from it what was relevant, and presenting this analyzed data to the operator.

Since TMI-2 great progress has been made in providing operators with relevant, analyzed data. Indeed, just as medical doctors can now receive diagnoses on-line on the basis of the symptoms they input, so reactor operators can now receive advice on how to handle crises on the basis of information put into their Artificial Intelligence control system. What we see here are examples of how prior intervention by sophisticated diagnosticians--either medical internists or reactor engineers--allow us to respond to crises intelligently; indeed, the sophisticated application of Artificial Intelligence begins to approximate some of the functions originally conceived for the Information Analysis Center. I do not claim that many of the on-line databases are yet able to provide analyzed information that is useful in handling a time-constrained event. But I believe we are moving in that direction, in any event the ability to call up data, even unanalyzed data, very quickly in general is helpful when a quick response is needed, whether in medicine or in reactor operations.

On the other hand, in research that requires the deepest kind of cogitation, and that is not constrained by time, I am less than certain that having all data at one's immediate fingertips is always best. After all, the two most important discoveries of modern physics, Relativity, and Quantum Mechanics, were made long before on-line information systems were dreamed of. Could there be a kind of Gresham's Law: that if information comes too easily, then researchers spend too much time absorbing the information, not enough time analyzing and contemplating the significance of the information? Could our modern very quick response system be imposing on our scientific enterprise a sort of journalistic flavor that is antithetical to the deepest understanding of what is going on?

What I say must be heresy: that if information systems provide data at too fast a rate, then the information process itself will tend to block out the process of understanding and analyzing. Yet I should think this possibility must have occurred to many in the information community. Our plea for information analysis centers in a way was a response, 25 years ago to our sense that something like this is a danger, or at least a consideration that we ought not to ignore.

Shortly after SGI appeared, I speculated on a possible role that information analysis centers might play in helping to codify, and organize, scientific knowledge and even to develop new, high order scientific insights. My example was the development of the shell theory of nuclear structure. Ever since World War II, data on nuclear properties--cross-sections, dipole moments, energy levels and the like have been accumulating at an astonishing rate. To find systematic regularities in this mass of data was not easy--yet, by the late '50s, the nuclear information scientist, K. Way, with her encyclopedic command of the data, was able to discern periodicities in neutron cross-sections as mass number increased. Helped by this inductive insight, Goeppert-Mayer and Jensen were able to devise the so-called shell picture of the nucleus; this model gave nuclear structure a coherence somewhat comparable to the coherence given to chemical structure by Mendeleef's Periodic Chart. I would say that in this instance the contribution of the information scientist, K. Way, supported by her information center, was of signal importance.

Can we see an analogous development in biomedical sciences? I have in mind the project to map the human genome. As I understand the matter, the complete map would involve some 10^9 base pairs. At the rate of 10^6 base pairs per day, the entire project will require some three years (and several billion dollars). Obviously, the codification of the information, perhaps more important, its analysis, will obviously be as important and as challenging a part of the genome project as the identification of the bases in the first place. The biomedical community will be justified in asking what do we do with the complete genetic map. Just as the immense body of nuclear data was eventually reduced to some kind of scientific order by systematic organization of the data, I suppose it is fair to suggest, or perhaps to hope, that information scientists, perhaps even information centers, will play an analogous role in helping make sense of the billion odd base pairs amassed by the genome project.

The genome project is hardly time-constrained; nevertheless its sheer magnitude, its enormous demands for data handling will surely place severe demands on those who will eventually try to make sense of the project. I should think that along with newer, ever cleverer and sophisticated computers, the project will require major data analysis centers devoted to this project. And, insofar as the genome project is not time-constrained, I would imagine the information handling system will have less in common with the on-line data bases and more in common with the kind of information analysis centers originally contemplated in SGI. I realize there are speculations; yet it seems clear that the genome project will prove a major possibly unprecedented challenge to the National Library of Medicine.

Let me close on a different note. As I view SGI from the perspective of 25 years later, I realize that SGI's insistence on the specialized information center as being key to the resolution of the information problem was rather naive. History instead has shown that what prevails is not what a self-appointed committee of savants conceives, but rather what the users--the working scientists--perceive as fulfilling their needs, and more and more, are willing to pay for. And in the free market competition of information tools, what have emerged as dominant elements have been a variety of tools--on-line data bases, review journals, new methods of electronic communication, libraries that take over some of the functions of information centers in providing information as well as documents--and where the users feel the need, information centers themselves.

Nor is this likely to change. As I pointed out earlier, John Tukey insisted that we "decrisify" SGI--because information is not a problem that admits of a single neat solution, but remains and always will remain a process that we must cope with. As Peter Medawar said in his 1969 address to the British Association, "We cannot point to a single definitive solution of anyone of the problems that confront us--political, economic, social, or moral, that is, having to do with the conduct of life. We are still beginners and for that reason may hope to improve. . . . The great thing about the race was to be in it, a contestant in the attempt to make the world a better place."

So it is with the information problem. Twenty-five years from now, on the 50th anniversary of SGI, I expect that the Leiter lecturer will return to these same themes and that the answer will be the same: that scientific information remains a serious issue, despite the existence of computers that make today's Crays look like toys; that the distinction between time-constrained and time-unconstrained tools will remain; and perhaps most important, that the National Library of Medicine will, as Peter Medawar implies, continue to be a major contestant in the race to maintain a viable scientific information system.