Researchers make little or no use of formal information systems. This is due, not so much to 'information apathy' on their part, as to faults inherent in information services as a whole. A formal information service may not come through at all. It is not capable of reaching behind a request to the real question that motivated it. By not being able to formulate answers in a way necessary to the researcher's logical momentum, the information service slows him down and may bring him to a standstill. All too often, a researcher's colleagues will fulfil these conditions--they tell him exactly what he needs to know to solve a problem, because they can grasp his problem and translate their answers in terms of that problem. The solution proposed is an information specialist who is a link between a research effort and a mechanized information service. He must be part of the research team as a senior member--so that he may know all about the work being done--and he must understand how the research product is to be judged and by whom. It is predicted that this approach will accelerate research work as much as six times, and that this will be the trend in years to come.
AS WE MAY THINK, INFORMATION SYSTEMS DO NOT*

William J. Paisley
Stanford University
August 1968

*With apology, if necessary, to Vannevar Bush, whose famous article, "As We May Think," in the July 1945 Atlantic Monthly, was one of the earliest and most significant manifestoes of information science.
I. White Elephants and Phantom Systems

-- Hans Selye reports of his information retrieval experiment that one user asked for "everything on endocrinology and stress." Urged to be more specific, he came back with "the effect of surgical stress on the hypophysis in rats." (1967, p.7)

-- Robert Taylor quotes a reference librarian: "The levels of frustration in using libraries are awfully high for most people. ...you are conditioned to feeling that the library is a place you almost have to drag something out of." (1967, p.3)

-- Harold Borko and Lauren Doyle observe: "All too often, workers in the documentation field have assumed a definition of the problem, in terms of particular user requirements, and have gone ahead only to build a 'white elephant.' ... Some of these systems rate very high in 'showpiece value,' but are so under-used as to constitute a severe embarrassment for the original system planners." (1964, p.7)

-- Alan Rees says it very succinctly: "busy people spending large sums of money, designing -- or attempting to design -- phantom systems for non-existent people in hypothetical situations with unknown needs." (1963, p.17)

Hans Selye's user knows the frustration of dragging information from a library. He doesn't really want "everything on endocrinology and stress," but he hopes that a request for everything might bring something relevant to his study of the rat's hypophysis, or pituitary gland.
There have been a hundred or so careful studies of scientific information use. They report various information-processing behaviors and attitudes of physical scientists as well as social scientists; scientists in universities, in industries, and in government labs; basic scientists as well as applied scientists; scientists in other countries as well as in the United States. Collectively, I regret to say, these studies return a vote of no confidence in libraries and technical information centers.

There is no time to summarize the evidence here, and I have reviewed it more than once elsewhere (1965, 1968), but study after study shows that information services -- from large federally supported clearinghouses to small company libraries -- are relatively ignored by the researchers for whom they were established. As examples (although, in fact, they are among the more successful operations) I would name the Defense Documentation Center and the Educational Resources Information Center, or ERIC, with which I am affiliated.

Four points need to be clarified before I move on. First, what do I mean by an information service being "relatively ignored?" I mean that the majority of a sample of potential users has never called upon the service, while many sometime users have tried the service just once.

Second, figures show that these services are serving someone: they disseminate millions of units of microfiche and hardcopy each year. How can a "relatively ignored" service be called upon for millions of units of information? One answer is that the pool of potential users in each case is very large, and requests from even a small percentage of users create a high volume of units disseminated. A more important fact is that institutional orders are being filled. University libraries and especially industrial libraries place standing orders and receive thousands of units. The microfiche and hardcopy flow from the shelves of one archive to the shelves of another, but this does not mean that researchers at those institutions requested the service or even know that a local collection exists. It's as if Ford's new cars were being sold to General Motors.

Third, data collected from the individual researcher, showing that he rarely uses a service, are irrelevant from one policy viewpoint: the archival function of these services must continue. Even if a year passed in which no one visited the Library of Congress, we would still need LC as an archive. If a year passes in which no one calls upon the Defense
Documentation Center, we will still need DDC (or its equivalent). Give up DDC, ERIC, etc., and we give up billions of dollars of information not available elsewhere, especially information contained in the "fugitive literature" of unpublished technical reports.

Fourth, past use of any information service is a poor predictor of future use. The time base of the "information explosion" is very short. Big Science is less than three decades old. Physics and chemistry made the irreversible step from little science to Big Science during the Second World War, followed in the post-war period by the life sciences and only now by the behavioral sciences. Information services were created as an afterthought. The federally supported information services are young and awkward. They perform zealously, if not well.

With much more behavioral research on researchers' habits and preferences in processing information, the scores of Big Science information services will become more responsive in meeting a researcher's information needs on his own terms. Then, I have every reason to believe, researchers will find the cost/payoff ratio in using such services more favorable and will call upon them more often.

But let me return to my theme: the data tell us that a researcher typically gets along with information stored in his head, together with information he can obtain from his colleagues and from a few much-used books on his shelf. Once in a while, to be sure, he conducts a "literature search," by which he means reviewing ten years of an abstracting service such as Psychological Abstracts and scanning the journals on his own and colleagues' shelves. However, he rarely goes beyond local resources (except in personal communication with researchers elsewhere, which is a different, and very significant, matter). He almost never consults the National Referral Center, the Science Information Exchange, University Microfilms, etc.

Some investigators of scientific information use (myself included) have inferred an "information apathy" from these data -- that is, we have inferred that the researcher just doesn't care whether he obtains information or not. I now believe that "information apathy" is the wrong inference. The situation is more complicated than that.
II. What Use Is Information?

-- In his Playboy short story, "Speed Trap," Frederick Pohl enters the mind of a harried, jet-age physicist to reflect, "You know, there really is no more stupid way of communicating information than flying 3000 miles to sit on a gilt chair in a hotel ballroom and listen to 25 people read papers at you. Twenty-three of the papers you don't care about anyway, and the 24th you can't understand because the speaker has a bad accent and, anyway, he's rushing it because he's under time pressure to catch his plane to the next conference, and that one single 25th paper has cost you four days, including travel time...." (1967, p.159)

What makes the 25th paper worth all that trouble? What use is information, anyway? How can we reconcile the notion of "information apathy" with the great time and effort researchers invest in some kinds of information? In the vernacular, what's so good about meetings and conversations? What's so bad about libraries?

My own thinking leads to five propositions:

(1) Processing information is more than a necessary evil in science. It is the central scientific activity.

(2) Collecting fresh data and searching results of past studies are complementary information-processing strategies. Up to a point, the more a researcher does of one, the less he needs to do of the other.

(3) If we define searching broadly enough to include conversations with colleagues and what Menzel (1966) calls "accidental" information gain, then these two strategies provide all the resource, all the raw material, that a researcher has to work with.

(4) The researcher's task is to synthesize information and to focus it on a problem or objective. If he is an applied researcher, the information will become new substances, devices, or processes. If he is a basic researcher, the information will become new principles of order and effect among phenomena,
for purposes of explanation, prediction, and control.

(5) In either case, he is judged by his product, not by the sheer amount of information he takes in.

There are two points to be made about the researcher being judged by his product. First is the question of who judges. Judgments of ingenuity, creativity, elegance, and overall success can be made by anyone who understands what the researcher has done. However, judgments made by those capable of rewarding him are of more concern to the researcher and to us. The applied researcher hopes that his product will be approved by his co-workers and supervisors, and perhaps by the production or marketing staff as well. Broader recognition of his accomplishment is unlikely, insofar as industrial secrecy and the anonymity of team research make it difficult for him to "stake a claim" through publication, copyright, or patent (there are exceptions, of course).

The basic researcher hears a different drummer. He wants to be judged by his peers -- that is, other basic researchers who are expert in his specialty. He expects that those who distribute other rewards, such as money or position, will defer to the judgment of his peers instead of attempting, as "laymen," to judge for themselves what he has accomplished. Amazingly, as Daniel Greenberg observes in his recent book, The Politics of Pure Science, basic researchers have managed in the era of Big Science to control billions of dollars of federal funds through a proposal-reviewing system that accepts peer-group (which often means in-group) judgment.

If information is raw material for a product, and if a researcher's performance is judged by that product, then the relevance and usefulness of information to him may be largely a function of who judges. It should not surprise us that information coming to the applied researcher from his co-workers and supervisors, or to the basic researcher from his colleagues, is so readily accepted. To some extent, such information already bears the endorsement of those who will judge the product to which it contributes.

Second, research so greatly transforms its raw material that both the quantity and the quality of the raw material can be obscured in the product. To put it more simply, the researcher knows that, if he wants to, he can fudge a bit on his raw material. How true this is of fresh data collection: if one experiment fails, pass it off and try another. Give just a paragraph
to the experiment that failed in reporting the one that succeeded. Fudging is even easier in a review of past studies. Failure to mention precedent studies will be interpreted variously as a negative judgment of their quality, or an assumption that the reader is completely familiar with them, or a concession to brevity, etc.

Even genuine ignorance of precedent studies, especially recent ones in contrast to the classics of the field, is judged tolerantly, since everyone knows that "the literature" is too disorganized to be searched exhaustively. In applied research the near-impossibility of finding highly specific, up-to-date information on substances, devices, or processes leads to a do-it-yourself norm in which many problems are taken directly to the laboratory without a literature search.

These circumstances inspire what I would call "information nonchalance" rather than "information apathy." The researcher knows that he is judged by his product, and he can vary the proportions of fresh data collection, literature search, conversation, and other information-processing activities to arrive at the same product. If literature search seems unpromising, he can increase the proportion of personal communication. According to 20 years of scientific information use studies, this is a modal pattern.

Let's be clear, however, that the researcher is nonchalant about particular sources of information, not about his need for information from some source (as "information apathy" would imply). He is like the motorist on a trip who stops indifferently at the first available gas station, the first available restaurant, and the first available motel. The motorist's nonchalance disappears when there is no available gas station, restaurant, or motel.

III. What's Wrong with Information Services?

-- John O'Connor discusses the question behind the information request: "A requestor who asks, 'How can Lissajous figures be generated by digital computer?' may be interested in a description
of how they can be generated more exactly at comparable cost by analog computer. Or a user who asks, 'What procedure can be used for preparing impurity-free ferric ethylate?' may be interested in a document naming a commercial chemical supplier of pure ferric ethylate.'' (1967, p.167)

-- In his study of medical researchers' information-seeking behavior, David Werner came upon one researcher who wanted to know whether substance X could be treated by process Y and was disposed to run the experiment for himself. "Dr. B indicated that, with the present state of knowledge, repetition would be less costly than retrieval." Actually, Mrs. A, an information specialist assigned to Dr. B's group, found the information in only 36 minutes, while Dr. B's technician took 4 hours to do the experiment over again. (1965, pp.94-95)

In Werner's interesting comparison of information search versus experiment, the library seems to have won by a margin of 3 hours and 24 minutes. However, we learn that Mrs. A's five steps in acquiring the information soon took her away from the library. In order, (1) she checked the John Crerar Library card catalog, with no luck; (2) she perused the John Crerar shelves, still with no luck; (3) she "called a friend at a chemical company who was familiar with processes related to Y," who referred her to a Dr. C; (4) she "called Dr. C, but spoke with Dr. D," who recommended Mr. P at a pharmaceutical house; (5) she "called Mr. P at the pharmaceutical house," who said that X could indeed be treated by process Y, and gave the name of two pharmaceutical houses which had further information.

The first two steps, both dead ends, took 25 minutes. The final three steps led her to the answer in 11 minutes. Unless we regard Mrs. A's friends and their friends as a composite "information service," by no stretch of the imagination can we say that an information service outperformed the laboratory in this comparison. Werner's study tells us three things: (1) the excellent, highly respected John Crerar Library was no help to Mrs. A; (2) Dr. B's insistence on repeating the experiment established an information "time value"
of 4 hours; (3) Mrs. A's network of friends and their friends came through with 4 hours' worth of information in just 11 minutes.

What's wrong with a formal information service is that often it doesn't come through at all.

What else is wrong? As John O'Connor observes, an information service answers, if anything, the request put to it, not the real question behind the request. In all the formal information services I know of, a question such as "How do I prepare pure ferric ethylate?" would be answered with a recommended procedure. The researcher's co-workers or colleagues, in contrast, would answer in the context of his effort and the quality of his product. They might recommend a procedure, or then again they might say, "Why bother to prepare pure ferric ethylate yourself, since the ABC Company supplies it locally at low cost?"

Serious as they are, these are the superficial defects of an information service. Of course a service can fail, and of course it can answer the wrong question. To make information services truly responsive, the essential defect we'll have to correct is a mismatch between information classification systems outside the researcher's head and "cognitive style" inside.

I won't attempt a full explication of cognitive style here, but I'm thinking of a certain structure for organizing knowledge, together with a certain strategy for solving problems, both structure and strategy exhibiting continuity in the person over time.

We know a great deal about information classification systems. We know the epistemological roots of the Dewey Decimal Classification, the Library of Congress system, faceted classifications, thesaurus-based systems, title-keyword systems, systems involving content analysis of text, and even systems based on bibliographic coupling. Without exception, these systems are proposed and defended on logical, rather than psychological, grounds. As far as I know, no system for classifying information has ever been proposed to complement the cognitive styles of users. Associative indexing (for example, Doyle, 1962) is as close as we come, and this system does not really derive from associative principles of verbal learning and meaning in psychology.

There are at least two reasons why users are expected to adapt to information systems rather than vice versa, even when a typical adaptation is simply to ignore the system. First, we know very little about knowledge
organization in the head. We have models and theories, for example the Gestalt tradition with such modern spin-offs as Boulding's *The Image* (1956). We have Tolman's cognitive maps (1948), Kelly's personal constructs (1955), Hunt's attribute dimensions (1962), and the common-sense psychology of thinking proposed by Bruner, Goodnow, and Austin (1956). We have linguistic approaches, for example Brown (1958). We even have measures in search of theories, such as Scott's categorization test (1962), which shows both the number and the nature of attribute dimensions a person uses to organize a small domain of knowledge. There are, of course, many other approaches, both conceptual and operational, but no answers yet to such practical questions as the difference in the way a basic medical researcher and a practitioner would cognitively structure a set of pathological conditions.

Second, we suspect that individual differences in cognitive style are the rule rather than the exception. Some work we have done at Stanford on the organization of knowledge with information-science objectives (primarily Coombs, 1968, with minor evidence reported in Paisley, 1967) reveals a great range and also a degree of individual consistency in such measures as cognitive complexity -- that is, the number of attribute dimensions used by a subject in a classification task.

Individual differences are no surprise. They are pervasive in the literature of information-processing and problem-solving. Our dilemma, however, is that we can adapt an information system to please one user most of the time, or to please most users some of the time, but not -- with 1968 technology -- to please most of the users most of the time.

The essential defect, then, is a mismatch between the normative logic, such as it is, of all information systems and the psycho-logic of each researcher's cognitive style. You could argue that I make too much of the mismatch, since bright people like researchers are extremely adaptable and, at worst, are only slowed down by an obtuse system. I would answer, first, that slowing down a researcher is bad policy (and probably has worse consequences than appear on the surface, as I will discuss later under the heading of "cognitive momentum"). Second, the data say that a researcher who is slowed down tunes out. That is, the fact that information services are relatively ignored is inescapable. Rejected the explanation of "information apathy," I must fault the service itself.
IV. A Responsive Information Service --
Its Performance Viewed Abstractly

-- Frederick Pohl's physicist in "Speed Trap" continues, "The pity of it is that electronic information handling is so cheap and easy. ... There's (picture phone), facsimile, telemetry, remote-access computation, teletype. ... And go farther, too. ... strip down a taped voice message ... 400 words a minute instead of maybe 60 or 70 ... problem-solving approaches to discussions ... Leaving three-quarters of our time for -- what? Why, for work! For doing the things that we know we ought to do but can't find the time for. ... I honestly think that we can do four times as much work as we now do ... land on Mars in 5 years instead of 20, cure leukemia in 12 years instead of 50, and so on." (1967, p.200)

It is not simplistic to see the researcher as a problem solver. His research project consists of macro-problems and micro-problems, problems with funds and personnel, with data collection and analysis, with instruments and computers, with reporting and publication. Most macro of all problems is approval of the product by those whose approval counts.

It suits my cognitive style to say that the researcher has a Plan, in the sense that Miller, Galanter, and Pribram use the term in their Plans and the Structure of Behavior (1960). The researcher's Plan is first an estimation of the problem-solving steps his research project will require. The experience of each problem-solving step may modify the remainder of the Plan. Unless the project is completely cut-and-dried, there will be gaps in the Plan to be filled on the basis of information collected in previous steps. For example, there will be a gap in the sub-Plan of data collection until some information is obtained about precedent work and about the feasibility of various methods.

We can follow Miller, Galanter, and Pribram further. We can think of each problem-solving step in terms of their TOTE (test-operate-test-exit) model. That is, the researcher "has an image of a desired outcome" (1960, p.38),
setting the conditions for which he must test.

"Planning can be thought of as constructing a list of tests to perform." (1960, p.38) The researcher moves from problem to problem, from sub-Plan to sub-Plan. In each case, he compares actuality with expectation, present state with desired outcome. If there is a discrepancy, he "operates" on the environment under his control with a strategy calculated to reduce the discrepancy. Then he tests again and, if the discrepancy has disappeared or fallen within a range of tolerance, he "exits" to the next problem. If the discrepancy is still too great, he must "operate" again to reduce it. So the TOTE cycle for one problem leads to the TOTE cycle for another problem or loops back on itself until success or exasperation break the loop.

In research a large proportion of all TOTE cycles are information-gathering cycles. Many of the information-gathering cycles are looped, or unsuccessful, cycles. According to studies of scientific information use, information-gathering cycles involving formal information services are more likely to loop than cycles involving the researcher's colleagues. Indeed, my theme could be restated that information services keep researchers going around in cycles, and this should not be.

Now we come to "cognitive momentum," by which I mean a certain satisfactory, self-sustaining, somewhat headlong movement through each TOTE cycle and from one TOTE cycle to the next. We experience cognitive momentum during the early phases of many research projects, when ideas and their corollaries come to us easily and click into place. We lose cognitive momentum when the research project begins to need information from outside our heads and we suffer delays in obtaining it.

The loss of cognitive momentum is distressing in itself, but more distressing is the compound interest we must pay to recover it. That is, each time we are looped back through a TOTE cycle, we must reconstruct the set of conditions we were testing in the first place. We are all familiar with the difficulty of reviving a research project after other obligations have forced us to put it aside for a while. It seems alien; we can't pick up the thread of the Plan. I contend that the delays and frustrations we suffer daily throughout the course of a research project have the same effect, except that in recovering momentum we spin our wheels for hours instead of days. Changing my metaphor, I believe it takes a certain head of steam to move through a
research project without losing track of objectives and conditions. The delays we must brook in gathering information cause the boiler to cool and the steam pressure to fall. Then, even when the information comes, we must wait for the head of steam again.

I know of no evidence on this point, so let me conjecture that research supported by a truly responsive information system would move forward, not half again as fast or twice as fast, but four or six or ten times as fast. Hence the appropriateness, to my way of thinking, of Frederick Pohl's prediction at the beginning of this section. With optimal support both in collecting fresh data and in searching past research, probably we could cure leukemia in 12 years instead of 50.

"Optimal support" can be defined, almost tautologically, as "the information that a researcher needs, remapped from its original state onto conceptual dimensions that the researcher has chosen to structure his thinking, and provided to him very quickly, cheaply, and without distracting nonsense."

No existing information service comes within a mile of optimal support. However, the researcher's colleagues often provide optimal support for particular problems. To put it more simply, a researcher's colleagues often tell him exactly what he needs to know to solve a problem. They are able to do so because they happen to have information he lacks (just as he has information they lack) and because they can, in a moment's conversation, adapt the information to his need through modification of vocabulary, omission of extraneous detail, and selection of conceptual structure.

We all know from personal experience how useful colleagues can be in this role. A fascinating memoir of the current year -- James Watson's *The Double Helix* -- will serve as a sourcebook on this topic for information scientists as well as psychologists of science. In their off-again, on-again progress toward a satisfactory model of the DNA molecule, Watson and Crick found the impetus for most forward steps not in their own laboratory experiments and certainly not in journals or books, but in conversations with colleagues in Cambridge and London. The Watson memoir shows us a detective collecting clues from informants who individually possess only bits of the solution but collectively possess almost all of it.

There are only two drawbacks in the constant use of colleagues as information sources. First, colleagues have professional responsibilities of their own, and even the most rewarding exchanges of information for esteem must be curtailed
if the man is unable to attend to his own work. Second, colleagues are not omniscient; they can be stumped.

Thus we have an always-available resource -- an array of information services -- that performs poorly and a resource that must be rationed -- the expert knowledge of colleagues -- that performs very well indeed. It seems that a responsive information service would combine the best features of both.

V. A Responsive Information Service --

Only Human After All?

-- From David Werner's study, mentioned above: "The brevity of Mrs. A's search hinged upon a personal contact who was the beginning of the series of connections leading to the information. ... Had (Dr. B) searched unsuccessfully, he would have still had to repeat the experiment. Since he had the technician's services, he chose what seemed the best method. On the other hand, had Mrs. A been available at the time, he might have found the use of her services to be a better alternative." (Werner et al., 1965, p.4)

The missing element in information service is people. Mediators. Middlemen. We cannot abolish the archives -- if anything, we need them more as the doubling cycle of scientific knowledge moves from decades down to a handful of years. However, I think that our dreams for mechanizing the archives and making them truly responsive to researchers and other users are dreams for the next century. Certainly we should mechanize the archives -- they are already too big for manual access -- but I think we should envision information specialists in command of the machine and linking the researcher with all stored knowledge.
As I interpret it, recent research on relevance tells us that the information specialist we need is not a reference librarian who works for the library. He or she is a relatively senior member of a research team who has developed information skills as a specialty just as other members of the team develop statistical skills, computer skills, etc. My evidence for saying that a reference librarian will not do is the Rees and Schultz finding (1967) that expert judges (e.g., scientists themselves) reached different relevance judgments from those of less expert judges (e.g., medical librarians). Here again we see a mismatch between the researcher and the information system, in this case represented by a librarian. If we wanted to explain the mismatch, several variables from the Cuadra and Katter studies (1967), such as "implicit use orientations" or expertness itself, have a demonstrated relationship with differences in relevance judgments.

Hence it seems clear to me that the information specialist must belong to the research team, must be intimately familiar with the history of the research project, and must be able to apply the same criteria of relevance that other members of the team apply. Above all, the information specialist must understand how the research product is to be judged and by whom, because all information-processing decisions made throughout the project -- fresh data collection as well as information search -- anticipate the criteria of that judgment.

I believe that we will see such information specialists on research teams in coming years. Not attached to teams, and not consulting for teams (Allen, 1966, suggests some reasons why such "outsider" roles work out poorly), but central in the teams, with publication credits and with discretion over funds and personnel. It will take us a few years to get over the notion that information searching is research assistants' work, and it will take the government a few years to realize that big time-sharing computer information systems are necessary but not sufficient conditions for responsive information service. But, after these adjustments, information specialists will be with us.

Meanwhile, to quote a recommendation of a National Academy of Sciences committee, we have barely opened the cover of the book we must fill with "intensive studies of the relationship of different information inputs to
research activities and scientific outputs, designed to ascertain how research processes and productivity might be affected by changes in information services and forms." (NAS, 1967)

REFERENCES


Paisley, W.J. Mr. Mits and his classification problem: a view from the user's side of the street. Paper read at annual convention of American Documentation Institute, New York, October 1967.


