

HELP - WHERE IS INFORMATION GOING TO?

Peter Popper

February 1975

WP-75-13

Working Papers are not intended for distribution outside of IIASA, and are solely for discussion and information purposes. The views expressed are those of the author, and do not necessarily reflect those of IIASA.



There are some indications that the present recession is bringing with it changes in the methods of information dissemination: some of these were already in evidence some while ago and have merely been accelerated, some are newish developments, some may be improvements and some the reverse. Before looking at the causes and effects, a brief review of the history and trends is appropriate.

First, however, a definition: information transfer implies the passage of knowledge from A to B in all forms - verbal, visual (including the written or printed word or diagram), and sensory. The earliest form of information transfer was by word of mouth, long before man began to write (animals of course continue to communicate by sounds). Symbols, hieroglyphics, writing all followed later; it was only the coming of the printing press with the use of individual letters by Gutenberg in 1445 which brought any notable change. The first technical book, by Roberto Valturio of Rimini, and the first medical book, by an unknown Italian, were produced in 1472; printed arabic numerals in a mathematics book first appeared in 1489. The thirst for knowledge of that age was mirrored by the rapid spread of the printed word (the first information explosion): by 1500 some 1000 print shops had been established in Europe and produced some 35.000 titles, with a total number of copies estimated at about 10.000.000. Half of these are believed to have been bibles. The first scientific periodicals to appear were the Journal des Scavants (1663) and the Philosophical Transactions of the Royal Society (1665). Growth in this area was slow initially, with a mere 10 journals published by the middle of the 18th century, 100 by 1800, 1000 by 1850 and 10.000 by 1900. Estimates for the late 1960's range from 50.000 to twice this number, the higher figure possibly being the more reliable. This exponential rise in journal titles, however, is not the sole growth factor; virtually all journals have shown a considerable year-by-year increase in the number of pages produced and hence in the number of papers (or bits of information) published.

Publication in journals is only one means of disseminating knowledge; there are books, reports, patents and papers presented at conferences (which themselves have shown a four-fold increase in number during the fifties and sixties). A conservative estimate puts the total number of bits of information to which we are exposed annually at 20,000.000 in the field of science, technology and medicine alone. All indications are that primary dissemination has been growing over the last quarter of a century at least at an exponential rate of between 3.5 and 14.5 %, depending on the parameter used, the lowest rate applying to periodicals since their inception and the highest to the number of articles produced

by engineers (the growth rate for conferences is about 7.2 %).

It has long been recognized that to keep up to date in a given field requires knowledge of as many bits of information as possible. Thus even in the early 19th century, learned societies tried to serve their members by including in their own journals abstracts of papers appearing in others: the secondary service had been born, and a lusty baby it proved to be. From a slow start it grew rapidly at an estimated rate of 8 %, going through a variety of phases - encyclopaedic, subject, mission-oriented, all-embracing, specialist. Enough has been written elsewhere about the history and future of this field - world plan has succeeded world plan (devised, naturally, by the secondary service producer and not by the user) - not to need detailed discussion here. Suffice it to say that mechanization (computerization) has not been the hoped-for panacea, but rather has become an end in itself; the terminology of information science (a grand term coined for self-glorification of a profession that never was and merely failed) serves as a warning against the misuse of language, which after all is the main carrier of information.

The rapid expansion in all fields of knowledge has naturally had its impact on the knowledge industry - primary and indirectly secondary. More scientists were produced, who in turn produced more information which had to be digested and circulated, drawing more people into the field and so on. Learned societies failed to react quickly enough to this trend, but commercial publishers did. The former, wishing to preserve their high standards and operating on limited budgets even in the years of plenty, could not publish enough or quickly enough; no such inhibitions troubled commercial publishers once it had been shown that the market existed and could be manipulated easily. The customer was tricked into subscribing to a volume which he took to cover one year, but which instead consisted of a fixed number of pages, so that every few months he had to pay for more. Having made the initial mistake, and probably being a librarian (some 85% of all journals sold go to libraries), he would not admit this to his management; and once a title is on a library subscription list, cancellation is almost unheard of. Production costs for a periodical, furthermore, are relatively low; provided the publisher can persuade some eminent person to act as editor (usually without fee), who will then appoint an editorial board (also unpaid), scientists will queue up to publish (again without pay): the main and perhaps only costs are paper, printing and distribution. If the journal is sufficiently well established it may even impose a page charge on the author's institution for the privilege of publication. Where fees are offered to authors, they are usually ridiculous in terms of the effort required, so that

periodical publication up to quite recently was a money spinner.

Again, the development of cheap air travel (and the affluent society) had an effect on conferences; these could be arranged in exotic spots to attract more and more participants, so that certain places set up an industry to cater for the new demand. From occasions where personal contacts amongst researchers could be renewed or made, conferences turned into almost unmanageable jamborees (can one really benefit from a meeting of 4000 spread over six towns in a country where only 5 % of the participants can communicate with the inhabitants?) With sufficient skill it proved possible to spend at least half one's year travelling from meeting to meeting, soaking up and imparting information of great value. How right Koestler was to call these travellers the modern call-girls.

Then there are monographs and reference books. If a researcher has written and presented a sufficient number of papers (as he frequently must to obtain academic and other appointments), and often irrespective of the inherent novelty of the writings, he will ultimately feel the need to write a monograph or be asked to do so. At worst it will be a textbook for use in a developing country, at best a popular best-seller not related in number of copies sold to merit of content.

Perhaps the above paragraphs appear cynical: but the fact remains that the growth in bits of information is far in excess of the growth in the number of scientists (or originators of these bits). This shows either that inventiveness and novelty have increased remarkably, or that there is repetitiveness or a drop in standards or both. It appears fairly obvious however, that science as a whole has let itself be exploited by this commercialized trend. Where this has led is all too obvious: far from easing the problem of information transfer, the contamination from the dissemination of ever more bits of doubtful value was rapidly leading to a complete breakdown in communication between scholars to which all could have plugged in, and to the withdrawal of real communication into the invisible college-type networks. At a time when interdisciplinary and open communication was becoming more essential information became less accessible.

Much discussion, effort and money have gone into attempts to solve the problems caused by this self-induced dilemma. Many solutions have been proposed, but none has proved as effective as one not considered - the effect of economic and other shortages.

Over the past three to five years, the costs of producing the printed page have escalated at an alarming rate (25 % per year), reflecting the increase in material costs (paper) and labour charges. Postal charges too have risen sharply (affecting periodicals more than books). Travel costs have taken their toll of conferences (so far, more in number of attendees rather than of events). As long as these increases could be absorbed by the customer, all was well, but this is no longer true today. Library budgets have not kept up with the decrease in purchasing power, and in real terms have shown a reduction. Furthermore there has been a decrease in the number of customers due to mergers in industry and closures. There is strong evidence that the buying of new books by students is showing a dramatic decrease. Purchases by established researchers even of paperbacks are also suffering. The writing has been on the wall for some time, but we failed to recognize the signs: fewer titles available off the shelf in bookstores, as the bookseller could not afford the capital investment of speculative stock holding: so-called rationalization by publishers who refused to sell single copies to cut down on invoicing costs, etc: smaller print runs, and so on. Moreover, the need to consider the economics of publication has led to a dramatic acceleration of the trend toward in-house publishing. The volume of institutional literature produced by the cheapest methods has increased at the expense of conventional publication. Such literature includes reports, preprints of conference papers (which may or may be not bound together as Proceedings), and even monographs with print runs as small as 100. Leaving value judgements aside, the chances of the information content becoming freely available through conventional methods of dissemination are slim (which may be the underlying aim).

We are thus in a transition stage where the imperfect mechanism devised for information transfer, based on the traditional pattern of information generation, is having to cope with an increasing number of bits of information, a large part of which is not being put into the conventional channels.

Perhaps it is appropriate to look at the user, who, together with the generator and distributor, makes up the system of information transfer. The user requires and receives two types of information input - the ad hoc and the directed. By ad hoc is meant all the information acquired through exposure to information-carrying media - colleagues, casual reading of all types of material, wireless and television, and so forth. Whether the received input is wanted or not is immaterial - the recipient will operate a filter and trigger mechanism and consciously or subconsciously store or reject.

Matters are different where the directed input is concerned: this is actively sought by the user for a specific purpose, and the user imposes certain criteria (which are often not fully allowed for by the distribution). The first need is for the information to be timely, i.e. available when the user requires it. (At the lowest level, it is no use hearing at five past twelve that the only train to the required destination left at twelve). Secondly the information must be relevant, i.e. have a direct bearing on the requirement. (It is irrelevant to know that there used to be a second train an hour later). Thirdly the intellectual content of the message must match the intellectual need of the recipient as far as possible. (It can be assumed that the ordinary enquirer would not want to know the type, horsepower, etc. of the engine that will pull the train).

Whilst hitherto the second of these parameters has received the most study, with lip service being paid to the first, little if any work has gone into examination of the third, and of the implications this has for the whole system. The reason is probably that this involves definition of a variable which is inconsistent and "unconstant"; i.e. both the message and the recipient vary in the sense that the same message may have different levels of intellectual content for a group of recipients, whose requirement in turn may vary with time (and who is to measure, and by what means, intellectual content or need?) Nevertheless a serious attempt ought to be made in this direction. There is an obvious need to refine both filters and amplifiers in an information network from this point of view.

It is of course true that the most effective way of information transfer is through the invisible college route, followed by the gatekeeper role. Now invisible colleges are the result of personal contacts, exchange of papers and letters etc. but entry is not easy for the newcomer. The traditional first contact by writing for a reprint is a victim of the xerox copy, rising postal charges and long postal delays, so that an established member may not always be aware of the new contacts quickly enough. The gatekeeper on the other hand may become a victim of an efficiency drive, since his visible contribution to output in an organization may be small, or the budgetary restraints on libraries may deprive him of much of his source material.

We thus have a fairly confused and conflicting state: an increasing flow of unfiltered information absorbing finite resources both in transfer and receipt of the product, with the impact of technology being at best neutral, but probably negative. The problem is one of reconciling the best

of networking, and incorporating somewhere a value assessment of the information being transferred. It would be difficult enough to arrive at a logical solution if adequate data were available, but they are not. We know roughly the total bits of information generated or put into the system each year and we have some idea of the number of recipients, but we do not know the distribution of bits among the different means of transfer, the intellectual content level distribution of the bits, or the absorption (receptivity) needs or capacities of the recipients. We thus have to engage in an extensive system of data gathering or try to reach an intuitive and pragmatic solution.

The object of the present notes is to stimulate some of the readers to let me know their information intake habits, their methods of disseminating information (whether as source or node) and any other thoughts which might help in trying to arrive at a solution.