INFORMATION SCIENCE TODAY

S. SAMPATH*

(*Additional General Manager, Bharat Electronics Ltd., Bangalore*)

Received on June 20, 1977

ABSTRACT

The growth of knowledge over the years has been phenomenal. One often wonders whether original research is cheaper than hunting for information already published. The electronic computer, a new entrant into the field of information science, plays a crucial role in acquisition, storage and retrieval of data. The communication satellites have made the data transmission process easier, quicker and efficient. The new trends in communication technology —surface-wave media, digital signal process, encoders and the use of optics—have opened up new and exciting possibilities. It is hoped that by 1980 complete programmes responding to significantly complex situations might be available. Concern is expressed that the availability of information to all might make inroads into the privacy of individuals and nations, but as experiments by IBM proved these can be contained.

Key Words: Communication satellite, computer, information science.

1. THE KNOWLEDGE EXPLOSION

If knowledge is power, the communication of all the available knowledge to potential users anywhere in the world, at reasonable cost, is a matter of paramount importance for the welfare of the human race.

Man's capacity to communicate with other members of his species, evolving over hundreds of thousands of years, is marked by four distinct and revolutionary developments. The first of these was the evolution of the spoken word, the origin of which has been lost in the mists of history. The second was the emergence of written language; and this happened several thousand years ago. The third leap forward occurred with the invention of the movable type in the 15th Century. And the fourth is the Electronics Revolution surging around us.

Prior to the year 1500, Europe was producing books at the rate of 1,000 titles per year. It took a century to get 100,000 books written. By 1950, the rate had accelerated to 120,000 books per year. What took 100 years before now needed only 10 months. By 1970, the output of books,

* President, Indian Society for Information Science., c/o. Library, Indian Institute of Science, Bangalore 560 012.

on a world-scale, had reached a figure of 1,000 titles per day, and the rate continues to rise. The number of scientific journals and articles are now doubling every 15 years. On a global basis, scientific and technical literature is streaming out at the rate of 60 million pages in a year.

An American firm has reached the conclusion that, for a piece of research likely to cost less than $ 100,000, it will be simpler to repeat the research than to hunt out all the published information relating to the problem. A stage has been reached where the most knowledgeable amongst us finds it extremely difficult to keep pace with the deluge of new knowledge even in specialized fields of interest. An eminent Physicist says: "On the subject of K-mesons alone, to wade through all the papers that have been published has become an impossibility." A specialist in oceanography expresses his difficulty thus: "I really do not know the answer unless we declare a moratorium on publications for a period of ten years."

The Electronic Computer

The Electronic Computer which literally burst upon the world scene in 1950, with its unprecedented power for analysis and dissemination of extremely varied kinds of information, in large quantities and high speeds, plays a crucial role in the area of knowledge—acquisition, storage and retrieval. In combination with powerful analytical tools, it makes for an acceleration in the accumulation and distribution of information over a wide spectrum of interests. It changes profoundly the intellectual environment in which people work. It generates new ideas on man as an interacting part of large systems and on his physiology and ability to learn, remember and make decisions.

Computers of immense power, speed and reliability have been built and are in use. There are large systems possessing multi-programming and multi-processing features together with facilities for interactive working and teleprocessing. There are mini-computers of various types for dedicated applications. Micro-processor technology, which is only four years old, breaks new ground, by placing in the hands of the electronic engineer and his associates a tool more versatile and application-oriented than anything he has known before.

The dominant features of the new generation of computing devices on the horizon are likely to be the following:

— emphasis on communication and control and not on computation;
— output-generation within the response time of the recipient;

— hardware-software modularity facilitating expandability and updating;
— provision for direct-data acquisition and report-generation;
— simultaneity of diagnostics and system-working;
— availability of an efficient low-cost programme generation as a part of the system itself.

The following user expectations are anticipated:

— a single system catering to diverse types of users (*e.g.*, salesmen, lawyers, legislators); diverse user facilities (*e.g.*, planning, design and production); diverse data bases (different sources and levels of aggregation) and fast retrieval of information;
— capabilities in terms of human mental tasks, such as simple enquiry (static enquiry, briefing and exception detection); complex enquiry (diagnosis, planning and selection); modelling (evaluation and optimization); design (construction); and discovery (intelligence retrieval from texts);
— an appreciation that in a situation in which man is the source or destination of the information, speed and capacity are of minor importance and quality will be the chief concern. From machines, this would demand 'reliability'; and from human beings, high quality presentation of data in an assimilable form requiring a variety of input-output devices;
— high priority for the interests of human users. Even as computer systems acquire eyes, ears, voice, arms, legs and fingers—to be able to communicate more efficiently with their physical and informational environments—highlighting the deficiencies of human beings as intermediaries between computers and the world around them, there would be many applications calling for man to play an important role in the loop 'interactively'.

Expectations such as these call for systems with their capacities pushed to the limit—systems handling a mix of jobs with diverse storage and processing requirements and with differing calling rates and priorities. These pose great intellectual challenges to computer specialists and information scientists. There is no doubt that they will be adequately met in the years to come.

Communication Satellites

A new dimension came into the domain of telecommunication about 15 years ago through the newly realized possibilities of relaying through communication satellites. A measure of the quick progress achieved in this area is revealed in the capability of one of the INTELSAT satellites, launched in January 1971, which handles 6,000 two-day telephone circuits or 12 simultaneous colour television programmes or tens of thousands of teletype circuits or a judicious combination of all of these. This single device embodies a channel capacity exceeding that of all the world's submarine cables put together. Yet it marks only the beginning of a new vista of possibilities. Looking at the global scene it is safe to predict that, by the end of this decade, many potential uses of communication satellites will have become realities of the day: direct dialling between subscribers anywhere on earth; control of commercial air-traffic; medical diagnostic service on an international scale; and the transmission of facsimile copies of entire books or newspapers, from one country to another, in a matter of a few minutes.

Just about all the information stored in computers will be retrievable almost anywhere by almost anyone with access to a computer terminal. This calls for the creation of computer information facilities that will serve either as large data-banks or as brokers between information sources. One sees the onset of a data-utility industry analogous to the electric-utility industry, with a Central Processor handling information at low unit-cost.

New Trends in Communication Technology

In recent years there have been phenomenal developments both in respect of the mathematical tools essential to the development of communication equipment and systems as well as in the design of the communication networks. Initially only electronic engineers were involved in the game of intelligence transmission; but today a host of others including chemists, astrophysicists, logicians and physiologists besides physicists and mathematicians are needed to help in resolving the problems thrown up by the need for meaningful man–machine interaction.

Mathematical tools occupy a place of prime importance. The role played by the Fourier-Transform in the development of concepts in telecommunication is well known. Among the newer tools developed in recent years are: the high-speed Fourier-Transform algorithm, Hilbert spaces, Galois bodies and analytical signals.

In the domain of physical devices, new methods are being introduced all the time resulting in enhanced performance in terms of speed, lower power consumption and increased scale of integration. Higher power levels are being reached both by solid-state devices and microwave tube systems.

Newer transmission media are being developed. These include the vacuum of deep space, waveguides and optical fibres. Ultrasonics has for number of years been used for signal p-ocessing. Initially the media employed were volume-propagation media offering little flexibility; but, about ten years ago, surface-wave media were invented and this technique offers better performance for lower cost than conventional systems.

The transverse filter and speed spectrum modulation are typical examples of the effort being made to develop new devices and techniques and establish new application areas. Digital signal-processing is increasingly resorted to because it brings in more stable performance by virtue of the more precise computation that it embodies. Encoders, that are rapid and accurate, which radar engineers and measurement specialists have used for many years for filtering and recording purposes, are finding avenues of use in communication technology. The use of digital techniques facilitates the shift toward the use of millimetre as well as optical frequencies as the means of being able to transmit large blocks of information. There will be increasing use of optics not just for data transmission and acquisition but for data storage and even for data processing though the earlier promises remain yet to be fulfilled.

There is a thrust toward the reduction of redundancy existing in the acquired information to enable it to be transmitted over narrow bands. Notable application areas include video-telephony, facsimile transmission of typewritten documents, weather charts, etc. Methods based on associating 'form recognition' with syntex-analysis are being experimented upon. There have been setbacks as well as some significant advances in this area.

The evolution in telecommunication and data handling is becoming less subject to restraints stemming from inadequacies in the hardware itself. The real problem appears to be to determine what services people want and which of these will augment the quality of their lives across the full gamut of information science.

2. Characteristics of modern Information Systems

A modern information system would be characterized by the following attributes:

— a large collection of records in a machine-readable form;

— not constructed for a single identifiable purpose but intended for a number of possibly quite different users;
— availability to a large number of users who may use terminals connected to the system by communication terminals;
— a system supported by a set of computer programmes for querying the data base, retrieving information and printing reports. Such programmes will invariably be written in forms that will be familiar and convenient to prospective users.

To the many users the system will appear to be an intellect-amplifier offering the means of dealing with a large number of variables, a wide range of statistical methods and models and of testing many kinds of alternative hypotheses.

Lately there is a trend toward storing information on-line. Such systems for management-information are file-handling systems in which any one, without knowledge of a programming language, can design, create and update files and, from them, print reports in useful formats. In current versions of the system, the files are stored on magnetic discs and they are available in a matter of seconds from a key-board terminal. The system may be intended to serve special purposes such as inventory maintenance, production control, marketing and sales analysis, or form part of a total information framework used to manage a business enterprise, an institution or a hospital. Experience with on-line data-handling systems indicates both the advantages and disadvantages that they possess. Any attempt to collect data from many sources and make the various components consistent with one another and amenable to on-line computer processing invariably turns out to be more costly and time consuming than originally envisaged.

The importance of information-systems, as tools for planning and development, makes it certain that such systems will continue to grow. Systems on land resources, urban development and legislative action will be developed at many centres. Demographic systems containing aggregate data on people will continue to be set up as the prime tool for social scientists and social planning.

There is a growing concern that computerized information systems will bring about an erosion of privacy and of human rights. Until recently, the records kept on any individual contained data regarding one facet of his activities and were stored at a place with a definite functional responsi bility, *e.g.*, school and college records, hospital records, tax records, etc. There is an increasing trend to set up data banks with records about an individual all at one place. A great deal of detailed information about him becomes available and may form, sometimes unfairly, the basis of sanctions against him. No country as yet has evolved laws to regulate data-banks as they relate to privacy. There are proposals under consideration in the international councils but the situation bristles with difficulties. Several questions arise for example: what kind of data may be gathered and by whom; how is the correctness to be verified; who should be allowed access to the data; and how can security be maintained. On the technical aspects, notable advances have been made in terms of computer hardware to meet special requirements.

Cryptologic techniques have been suggested as a logical means of shielding a computer's store which might otherwise be readily retrieved from the machine Depending on the purpose of the data bank, some or all of its contents should be inaccessible and unreadable to anyone but an authorised user. Who is to be judged an authorised user would be a matter of public policy. The fashioning of the technical means is the province of the scientist and the engineer.

An experimental system, developed by IBM and called LUCIFER is an example of the success achieved in this regard. It comprises a file of passwords and a correlated file of cipher-keys within the computer. Access to the system is limited to those persons who can furnish a legitimate password at sign-on and who can furnish the system with a cipher-key that corresponds to the password given. The cipher-key selected at the computer and the one in LUCIFER must be identical. The strategy is designed to prevent disclosure of the information to anyone not possessing both the strictly private key as well as the individual key known to the computer.

III. The Scientific Basis of Information Storage and Retrieval

The field of information storage and retrieval uses a wide range of tools drawn from the abstract sciences and deals with more precise propositions than are met with in social sciences. But its fundamental measurements are still dependent on opinions and attitudes. In the case of text-retrieval, the relevant decision is clearly an opinion; and in the case of numerical material, the opinion comes through subtly.

Being in an area where logical considerations prevail over physical parameters, the discipline of information storage and retrieval derives its tools and techniques from the mathematical sciences. There has been notable, mutually beneficial interaction in algebric weighting procedures, logic, statistics and computational statistics. There are aspects where exact measurements are not feasible. Complex mathematical manipulations utilizing exact data have no superiority over simpler techniques using fundamental parameters. While mathematical tehniques are undoubtedly important, critical judgement will be the basis for formulating useful applications.

In the area of text-retrieval, the major development is the increasing availability of machine-readable texts. This makes it possible to introduce automatic indexing for certain types of data bases. Equally significant is the advent of solid system evaluation work. Systematic comparisons of system component functions help to streamline procedure for choices relating to document-length, index-language, etc. There has been relatively slow progress in the areas of statistical and linguistic analysis of natural language for meaning. There are no theoretical structures yet characterizing words and their interaction with context. Computers are not yet able to deal with meaning.

The study of numerical retrieval, particularly in relation to management information systems, is among the most fascinating areas of computer sciences. In the task of building up genuinely useful systems, designers face a tough optimization problem. Procedural systems employing primitive operators will prove to be too expensive in terms of machine and programming time. Systems with a wide range of special operators will demand large computer storage and are likely to confuse the programmer. The solution will have to be one that is responsive to the operational environment.

IV. The Social Consequences of the Communication Revolution

In the course of an address in which he spoke about 'the diverging human function,' Dr. Vikram Sarabhai posed this question:

> " How shall we preserve democracy, when the media of mass communication provide means of instantly reaching downward from centres of authority, and, short of public agitation, there is no authorized channel for the reverse feedback for controlling the governing system ? "

It is clear that, in today's society, men's consciousness will be shaped in a manner fundamentally different from the literate or pamphleteering mode of transmitting information characteristic of the industrial age. The trend would seem to be toward the aggregation of the individual support of millions of uncoordinated citizens, easily within the reach of charismatic personalities effectively exploiting the latest communication techniques to manipulate emotions and control reason. Reliance on television—and hence the tendency to replace language with imagery, with the latter unlimited by national frontiers—tends to create a highly impressionistic involvement in global affairs. There is the poignant story of a three-year old American child—an avid T.V. viewer and accustomed to the imagery—as she watched the funeral ceremony of former President Dwight D. Eisenhover, turning to her mother to ask "Mother, who shot him?"

Aristotle stated that the size of a political unit is determined by the range of a single man's voice. He was a citizen of the Greek World where all the residents of a City State could assemble at one place to transact their common business. By a similar yardstick, the entire world today is a potential single community. There is no part of it that is not within the range of a single man's voice. The significance of the service that electronics renders to information science is that, in the world of today, information has become a central economic resource.

The communication revolution affects not only the storage, retrieval and dissemination of information but it infringes upon the traditional concepts of national sovereignty as well. International postal communication can be blocked or censored; but the obstruction of the flow of information crossing national frontiers by radio-communication methods would require very complex and costly jamming equipment. Radio and television signals via communication satellites may be picked up by individual communities without their having to pass through national relay-networks. Here lies a potential means of political influence and control. Besides affecting the horizontal barrier of nationalism, the new communication capabilities will assail the vertical stratification of society. Information reaching out to remote regions on a scale not feasible before will create new aspirations and generate new demands among the people in these regions. In the Space Age, communication media will break through the traditional fragmentation of society and reveal outdated limitations, emphasising that human rights and social aspirations, knowledge and information can no longer remain the monopoly of rich nations or of the privileged class in a poor country.

5. DATA AND DECISION-MAKING

Way back in 1810, Wolfgang von Goethe wrote:

> "The modern age has a false sense of superiority because of the great mass of data at its disposal, but the valid criterion of distinction is rather the extent to which man knows how to form and master the material at his command."

The appropriateness and applicability of this statement to the conditions of today are beyond question. It sums up admirably our problem: We have vast quantities of data and make use of the latest means provided by technology. Yet we devote scant attention to the critical factor in decision-making, namely the intelligent use of information.

We increasingly encounter situations for which there are no precedents and decisions have to be made in those areas. This implies the emergence of new intellectual competence based not only on analytical and mathematical genius but organizational and inventional skills. We also need to cultivate people who will make decisions that were previously made by a scientist or a group of scientists. In a world which increasingly becomes one in which we may collect all the data that we think of, the premium will be on our ability to ask the right kind of question and wisely interpret the information flow patterns in order to be able to make value judgements worthy of the effort.

6. LANGUAGE AND HUMAN COMMUNICATION

In his illuminating essay on Language and Human Communication, Lewis Thomas says:

> "Ambiguity is an indispensable element for the transfer of information from one place to another where matters of real value are concerned. For meaning to come through there has to be a vague sense of strangeness and askewness. Speechless animals and cells cannot do this. When a bee is tracking sugar polarized light, observing the sun as though consulting its watch, it does not veer away to discover an exciting marvel of a flower. Only the human is designed to work this way, programmed to drift away in the presence of locked-on information, straying from each point in a hunt for a better, different point.

If it were not for the property of ambiguity, for the sensing of strangeness that words in all languages provide, we would have no means of recognizing the layers of counter-point in meaning; and we might all be spending our time sitting on fences staring into the sun. Undoubtedly we would be making everyday use of the alphabet and have acquired the capacity for small talk. But we would not have been able to evolve from words to Bach."

The attributes of ambiguity and strangeness in language used as the vehicle of human communication will form the ever receding frontier out in front in the cultivation of information science.

7. The Prospect of Thinking Machines

Natural language comprehension wherein a human being can converse with a computer in everyday English for instance has made good progress in recent years. There is reasonable expectancy that by 1980 we will have available computer programmes responding to significantly complex situations. These will come through improvements in computers memory systems, electronic speed, computational capacities and information-processing power. Capabilities already available have not been fully utilized yet.

Are we likely to see the development of machines that will succeed in amplifying our cerebral power. Jagjit Singh, in his essay The "Mystery of Human Wisdom," argues that computers will never be able to think because they lack autonomous desire or freewill and also because the thought-process is not amenable to the step-by-step counting routine which is the mainstay of the machine. We need a breakthrough in our understanding of the neurophysiological complexity of the human cerebral cortex; and it does not appear to be around the corner. There is a kind of indeterminacy that acts as a barrier—arising from the fact that, the more micro our neurological probe, the less macro is our comprehension of the overall working of the human brain.

8. Conclusion

In the words of Samuel Butler,

"Life will ever remain a pattern in which we will be obliged to draw workable conclusions from insufficient data."

It is the main concern of the discipline of information science to deal with this insufficiency. Let us hope and trust that, in spite of all that we do in this regard, a modicum of this inadequacy will persist, giving us mastery over the machine and machine-assisted systems, and that there will be continuing advances, in the realms of computers and communication facilities, to serve man in his quest for knowledge and information—not to denigrate him or diminish his place in the Universe.