A positive but undeluded view of ADP for intelligence reference services.

AUTOMATION FOR INFORMATION CONTROL

Paul A. Borel [1]

Last July the *Saturday Review*, in a special issue on the automation revolution, equated the computer with the atomic bomb as a technological development formidable enough to make a turning point in human history. Some months earlier a *Newsweek* report entitled "Good-by to Gutenberg" gave readers a glimpse of other things to come in the field of information technology: a photosensitive crystal the size of a sugar lump that is capable of containing images of 100,000 pages; a lensless photographic system which could lead to three-dimensional home television; a no-contact, no-pressure printing technique that can write on sand, print a message on a pizza, or put a trademark on a raw egg yolk. Marshall McLuhan, in *Understanding Media—the Extensions of Man*, predicts that books and newspapers will in time no longer exist, that publishing will give way to an active servicing of the human mind through research packages done to suit individual needs.

Spume and Substance

Over the last 20 years we have learned not to depend on such extravagant promises to get us out of our practical difficulties. I do not doubt that amazing developments will continue to take place. I do doubt that we can count on them for early and revolutionary solutions to our data-handling problems. The problems created by the exploding mass of intelligence information have a habit of staying well in front of innovations.

In CIA we are currently upgrading our computer facilities with third-generation hardware, infinitely superior to our initial gear. Yet the contribution of the computer to the task of producing intelligence

[1] Adapted from Section IV and Appendix C of the author's presentation to the London Intelligence Methods Conference, September 1966. The full paper, *Controlling Intelligence Information*, 51 pp., covers trends in information, their impact upon intelligence, and the controls used, including the use and promise of advanced information processing systems.

is still both specialized and very limited. We have secure telephone communications; but these are far from ideal, with few instruments and high costs. Great strides have been made in our printing establishment. Still, the lapse between preparation of copy and its availability to the reader can be measured in weeks rather than days for non-priority items. Reproduction techniques have shown major gains. But material received in such poor quality that it cannot be microfilmed runs in some categories as high as 20%. We have improved our means of instructing reporters. Yet 50% of the titles in some report series have to be rewritten to reflect the content properly.

This experience gives ground for caution against any wholesale abandonment of the workable (if less than satisfactory) old in favor of the glamorous but untried new. Nevertheless, this is a time of important new developments in practical means for information handling, and intelligence should pay more attention to what is going on in this field outside. As never before, we have opportunities to capitalize on the work and ingenuity of others to relieve some of our own problems. Much of the work done outside is solid and relevant. We ought to use it, pick-a-back, whenever we can.

Active State of the Art

Let me mention a few such outside developmental activities touching the library science field. Two programs are being carried out in the academic community at large. One, named Intrex, for "information transfer experiments," has been called a step toward a dial-a-thought world. It is setting up an experimental laboratory to test ways of giving professors and students instant access to information. Xerography, film projection, and telephone communication between computer and user are planned. Basically, the experiments will attempt, first, to automate and rationalize the functions of libraries and, second, to develop a computer-based information transfer network. Another program, under an organization called Educom, the Interuniversity Communications Council representing over 30 universities in 20 states, is evaluating the significance for higher education generally of electronic hardware (computers, light pens, graphic displays), and software (computer programs).

A number of individual university libraries have forward-looking programs, Washington State and Florida Atlantic to name only two. The latter has the distinction of being the first in the United States to have introduced data-processing methods and techniques into its

operations at its very beginning. Washington State, on the other hand, is converting from traditional library methods to a totally on-line system which offers multiple remote access to a single library record. Sharing the time of the university computer (an IBM 360/67), it will be able to reduce typing substantially, eliminate duplicate manual files, and give complete control of each item's location and status in the library.

The value of these projects to us is that they are comparable in size to those which intelligence libraries may undertake. While much valuable information has been published about ways to automate the Library of Congress, the sheer size of its holdings makes many of the parameters of that undertaking inapplicable for us.

The publications of professional engineers, documentation specialists, and experts in various aspects of the information-handling industry are also increasingly solid and relevant. Some particularly useful books and articles are listed in the bibliography. These make evident my point that outside the intelligence community there is much wisdom and talent which we have neither tapped sufficiently nor used effectively because we are ill organized to do so.

CIA Applications

Certainly CIA has had in the past no organization worthy of the name to identify this outside work and relate it to our own improvement programs. This gap has now been filled with the organization of our Intelligence Sciences Laboratory, which is acquiring its own computer and associated equipment to provide an experimental environment closely approximating actual operations. Illustrative of its prospective areas of activity are on-line analytic processing, pattern recognition, language and text processing, and speech and audio manipulation. We will thus better bridge the work done outside and our own EDP-related operations.

While CIA pioneered much automatic information processing with its punched-card equipment, our experience with general-purpose computer operations is short of six years. In those years we have very considerably expanded our use of these machines. We now have a major computer center providing counterintelligence and operational support, another serving intelligence production, and a third devoted to imagery analysis. In the last three years we have reached the point where we see computer applications in almost every element of the intelligence cycle, from the management of collection requirements to the printing of finished intelligence.

Some of these applications are initiated by the analyst charged with intelligence production. Where the task data is numeric (like military-economic costing or agricultural production and soil moisture statistics) or in simple standard format and the data preparation is done by the EDP staffs, such applications, especially in the economic area, have given high-yield products which require no great investment of effort from the analyst. A wide range of other applications of this type remain to be tried, including computer control of Soviet scientific and technical literature. Work of this kind, involving primarily collaboration between an analyst and an applications expert from our central computer facility, can be characterized as special projects.

A much more ambitious application is our current attempt to change fundamentally the present method of doing business in the Agency's Office of Central Reference—our Project Chive. In one degree or another this general project will change the way hundreds of analysts are now working.

The Chive Project

The need for Chive arose from developments over the past nearly 20 years, during which we evolved a number of special reference services to support the production analyst. The multiplicity of classifications, of indexing tools used for control, and of formats employed in collection, dissemination, storage, and retrieval made it increasingly difficult to meet customer needs. The problem of heterogeneity was compounded by the increase in volume of data received and, with the passage of time, the volume in file. Moreover, the intelligence production expected of the analyst today is characterized by greater sophistication and shorter deadlines. Project Chive is designed to help him meet that challenge.

There are many unique features of the Chive approach. The project team is an integrated group drawn from production, reference, and computer components of the Agency, and it includes contract personnel as well. Experienced operators of our information systems have been given training in advanced techniques and placed in charge. The prospective user of the system is drawn in as active participant. A single important geographic area, China, has been selected for first application of the system, and even here it will be conducted as a pilot operation in parallel with the old system to permit experimentation before it carries the whole load.

From the user's point of view, the system should provide a number of advantages, most immediately:

All-source retrieval from a file system covering every type of printed document, including maps and photos, at whatever classification level.

Single-point retrieval service organized by geographic area, as opposed to the old multiplicity of indices and registers.

Literature searches that turn up all books, documents, reports, etc., that bear on a subject in question.

Information searches that turn up facts in answer to specific questions, facts concerning foreign personalities, organizations, installations, and activities.

Counts, whether of Algerian students in the USSR, public appearances of the Chinese leaders, or CIA intelligence reports on Haiti, and the trend of changes in such counts.

Detection of redundancies and inconsistencies in the system store.

Less matter-of-course but not at all visionary are services like the following:

Automatic inference-making—manipulating the wide variety of stored facts about people, institutions, and activities to produce new hypotheses about their character and connections. The variety of problems to which such a capability might be applied would depend on the ingenuity of the intelligence analyst.

Machine-assisted language translation, both machine translation of Russian documents and machine conversion of oral translations from other languages into printed documents in English.

Analyst referral service from a central directory or "profile" of human sources with expert knowledge in special subjects.

Remote querying which will enable users to interrogate and maintain from their own offices special-purpose files in the central system.

Half a Loaf

We see the development of this improved system as extending over ten years and many difficulties. Only because computer technology, capability, and capacity are what they are today and will be tomorrow do we dare count upon the success of the project.

Even so, there are risks. In this costly field you try to reduce the risks, but after you have done all you can they are still considerable.

We recently had a series of meetings with Dick Brandon of Brandon Applied Systems during which he reviewed with us the experience of others in using computers. He said that of 16,000 installations equipped with 27,000 systems in the United States today, 40% are unsuccessful in the use of their computers. This means that 6,500 organizations are not deriving economic benefit from them or are not achieving their objectives. In 90% of these cases, schedules and budgets have been exceeded. The main reason for this, in Mr. Brandon's view, is that the people using the machines are 'way behind the technology. They are not capable of utilizing the machines' capability.

My own rule of thumb in the application of machines to non-numeric problems is this: expect half as much in twice the time at twice the cost. If you get it you can count yourself lucky.

BIBLIOGRAPHY

Books

Planning Conference on Information Transfer Experiments, Massachusetts Institute of Technology, 1965. *Intrex: Report of a Planning Conference on Information Transfer Experiments, 3 September 1965.* Edited by Carl F. Overhage and R. Joyce Harman. Cambridge, Massachusetts: MIT Press, 1965. 276 pp.

King, Gilbert W., and others. *Automation and the Library of Congress.* Washington: Library of Congress, 1963. 88 pp.

Licklider, J. C. R. *Libraries of the Future.* Cambridge: MIT Press, 1965. 219 pp.

Bourne, Charles P. *Methods of Information Handling.* New York: J. Wiley, 1963. 241 pp.

Greenberger, Martin (Ed.). *Computers and the World of the Future.* Cambridge: MIT Press, 1962. 340 pp.

Feigenbaum, Edward Albert, and Julian Feldman (Eds.). *Computers and Thought.* New York: McGraw-Hill, 1963. 535 pp.

Borko, Harold (Ed.). *Computer Applications in the Behavioral Sciences.* Englewood Cliffs, New Jersey: Prentice-Hall, 1962. 633 pp.

U.S. National Science Foundation, Office of Science Information Service. *Current Research and Development in Scientific Documentation.* No. 14. Washington, 1966. 662 pp.

Conference on Libraries and Automation, Airlie Foundation, 1963. *Libraries and Automation; Proceedings,* edited by Barbara Evans Markuson. Washington: Library of Congress, 1964. 268 pp.

U.S. National Academy of Sciences, National Research Council, Automatic Language Processing Advisory Committee. *Language and Machines: Computers in Translation and Linguistics.* Washington, October 1966.

American Documentation Institute. *Annual Review of Information Science and Technology,* Vol. I, edited by Carlos A. Cuadra. New York: Interscience Publishers, John Wiley & Sons, 1966. 389 pp.

Articles

Fiock, L. R., Jr. "Seven Deadly Sins in EDP." In *Harvard Business Review* pp. 88-96, May/June 1962.

Flood, Merrill M. "The Systems Approach to Library Planning." In *Library Quarterly* 34:326-338, October 1964.

Knox, William T. "The New Look in Information Systems," Washington: Office of Science and Technology, Executive Office of the President, 1966. 18 pp.

Salton, Gerard. "Progress in Automatic Information Retrieval." Reprint from *IEEE Spectrum* vol. 2, no. 8:90-103, August 1965.

Shapiro, E. B. "An Evaluation of Computers in Text Editing." Menlo Park, California: Stanford Research Institute, 1966. 31 pp. (SRI project 5849).

Swanson, Don R. "On Improving Communications Among Scientists." In *Library Quarterly* 36:79-87, April 1966.

Schwartz, Judah L. "Computers and the Policymaking Community." Livermore, California: University of California, Lawrence Radiation Laboratory, 1966. 19 pp. (AEC contract no. W-7405-eng-48). (UCRL-14887; TID-4500, UC-32).

John McCarthy and others. "Information." In *Scientific American*, September 1966 (special issue), pp. 65-260.

U.S. Federal Council for Science and Technology, Committee on Scientific and Technical Information. "Recommendations for National Document Handling Systems in Science and Technology." Washington, 1965. 20 pp. Appendix A, "A Background Study," vol. 1. Washington, 1965.

"The New Computerized Age." In *Saturday Review*, pp. 15-37, 23 July 1966.