ARCHIVES AND AUTOMATION

By
J. SHAW,

Assistant Commissioner, ADP General Development Branch,
Australian Public Service Board

Introduction

It is reasonable to expect that the archives institution of the future will differ from that of the present to about the same extent that the present archives institution differs from the repository of clay tablets established by the King of Babylonia.

Primarily in this session I would like to set up a scenario for future archival systems and in doing so propose to briefly traverse:

* the background to technological advances relevant to archival activity;
* main areas of archival processing;
* processing strategies, problems and considerations associated with the introduction of automation into the archives;
* the possible future pattern of automation applicable to archival work.

A definition of the term automation is not offered; its meaning, hopefully, will emerge from the subsequent comments.

Background to Technology

Since the archivist is at the end of the line he is faced with accommodation to the technologies adopted by those providing him with his basic material. On this ground alone he must become familiar with advances in technology. Further, since his task of providing an extensive service is not an easy one, he will be further obliged to ascertain the extent to which current technology will aid him in his work.

As to basic archival material, after a relatively slow growth in the variety of forms of media over past centuries the range of methods of production has increased substantially over the last 50 years (combined with large expansions in volume). This situation obviously magnifies the problems of archival information storage and retrieval.

On the other hand, expectations as to timing of introduction of major advances in technology frequently have been extremely optimistic. Notwithstanding the over-optimistic estimates virtually in every case the technology advances forecast eventually become available. There seems to be a 10-year time lapse from the date predicted by the supplier for the introduction of a major innovation to the time the technology becomes embedded within society, with first full availability at the five year mid-point.

It is seen as essential, however, to keep a close watch on the forecasts being made (tempered by a recognition of suppliers penchant for over-optimism) in designing systems to ensure that later movement to newer technologies is not unduly retarded. In the archival context I would see the following as the expected major technological advances which need to be kept in mind:

	Expected Initial Availability	Full Availability	Wide Spread Implementation
Data Transmission Network	1970	1975	1980
Data Base Management Systems	1971	1976	1981
Mass Data Storage Systems	1975	1980	1985
Multi-font Type Readers	1980	1985	1990
Picture Transmission Network	1990	1995	2000

Experience has also shown that any major computer system has about a 10-year life cycle; but the replacement system generally steps off from the shoulders of that being discarded.

Areas of Archival Processing

The major areas of the archivist's functions to which automation might be applied appear to comprise the following:

* holdings;
* indexes;
* management and control

Holdings

Presumably present holdings are predominately in paper form. Questions arise whether, and the extent to which, such material might be transcribed to machine language for storage in a form suitable for computer processing. Answers, no doubt, will largely depend on:

* emotional, legal and logical reasons for retention in original form;
* advantages to be derived from retention in machine language form;
* costs of transcription to machine language.

Use of the computer offers opportunities for adoption of new modes of physical storage (e.g. the latest document received being allocated the next available storage position in place of retention in some other logical order – the computer keeping track of document location and available storage space). A large quantity of material in the future can be expected to be in other than paper form (microform, video tapes, magnetic tapes); there will be an increasing requirement for provision for storage and reading of, and extraction of information from, such media.

Indexes

There are a variety of ways in which information about holdings can be computer stored and computer accessed. Choice of mode or modes of indexing is a fundamental part of selection of the computer processing strategy.

Holdings Management and Control

Automatic data processing systems can be developed for such functions as holdings inventory recording and control; holdings usage recording and control; enquiry recording, answering and control; and production of statistics for archive management purposes.

Administrative Functions

In most archival organisations it would be expected that computer systems for administrative and similar purposes (e.g. personnel record and payroll processing, financial ledger keeping) would not be allocated a high priority, if to be introduced by the organisation, or those already developed and operated by other responsible authorities would be used.

Processing Strategies and Considerations

Movement to extensive use of computers for archival purposes will require the making of a number of significant decisions. The main decision points will lie in the following areas:

* the overall system strategies to be adopted in relation to holdings and indexing;
* stages to be adopted for conversion to extensive computer operations;
* plans for acquiring resources (especially manpower), equipment and funds.

System Strategies

Decisions on system strategies to be adopted must take into account the following aspects:

Holdings

* forms in which holdings are to be stored and accessed under a computer system (considerations include the extent to which information will be held in the form of either full text, abstracts, key words, titles or a combination thereof);
* the extent to which access must be restricted to preserve appropriate levels of confidentiality and privacy;

* the required speeds of response to enquiries;
* the geographical spread of points of access to the system;
* prospective interlinks and exchanges with other archival systems.

Management and Control

* nexus between the holdings/index systems and the management and control system (e.g. after a search yields the document number the inventory sub-system may need to be accessed for information on location of the document);
* nexus between sub-systems of the management and control system (e.g. the inventory sub-system may contain information as to the right of access to a document; the borrower sub-system may contain data on the borrower's access rights — cross-checking between the sub-systems would be necessary).

It is emphasised that the aspects of confidentiality and restriction of access assume significant importance with on-line computer systems since present procedures whereby staff act as vetting agents probably would not exist.

Development Stages and Plans

Stages of development adopted must obviously be feasible in the light of the technology then current and coming forward and the prospect of obtaining required resources.

The first phase of consideration of introduction of computers to the work of the archives would comprise a feasibility study. The object of the study would be to determine the extent to which automatic data processing could be applied either on the grounds of economy and/or of being the only means of satisfying future requirements. Subsequent phases would include those of strategic system design, detailed systems design, equipment selection and installation, system construction, system implementation and system maintenance. Experience has shown that for major computer systems there is an elapsed time between feasibility study and initial system implementation of the order of 40 to 50 months.

Processing Considerations

The following are major considerations which would need to be taken into account in the development of a computer system for archival work:

* the feasibility of use of the computer centre of another organisation either on a short term or long term basis;
* use of the general data transmission network (dedicated data transmission system could be expected to be too restrictive and probably too costly to introduce and operate);
* adoption of data transmission standards;
* selection of common type remote terminal units;
* employment of appropriate programme packages or standard computer information processing systems such as data base management systems and those for data retrieval.

If a decision were taken that all information received in the form of documents would be held within the system sometime in the future in machine language then early consideration would need to be given to converting all typewriters and computer printers of organisations forwarding documents to a standard machine readable type font.

Possible Future Pattern of Automation

A development pattern seen as technically feasible could comprise three stages with completion of introduction by 1980, 1990 and 2000 respectively. Whether any such conversion to computer working was justified and actual plans to be adopted would of course be determined by the feasibility study.

Stage I (1980) would be expected to comprise a computer system primarily designed to provide on-line index searches and responses. As well there could be on-line linkages with indexes of other organisations. The system would incorporate mechanisms for controlling access and for providing management with information required. While a growing quantity of holdings is likely to be in machine language form the system would not provide for machine scanning of the contents of such holdings as a normal process.

Stage II (1990) would primarily comprise the connection to the computer system of mass storage devices in which the holdings information (i.e. contents of documents) would be retained rather than the documents themselves. A single mass storage device by that time would be expected to be capable of storing some 10^{12} characters of information (i.e. the equivalent of one million bibles). By that time it would be expected the general data transmission network would have been upgraded to permit widespread and rapid transmission of bulk data.

The push-button telephone would also be expected to be in wide use by this time. Enquiries could be keyed into the central system with synthetic voice response; in practice, however, this facility could be found to be of only limited value in the archival context.

Stage III (2000). In view of the recent advances in the development of the optical fibre (a thin tube capable of carrying immense volumes of data in the form of lightwaves) and longer term plans for application to data transmission, the expectation is that before the end of this century a general network will have been established permitting data to be transmitted in picture form at high speeds. The likely prospect of this facility by the end of the century will require contemporary designers of archival systems to weigh up the advantages and disadvantages of conversion of present and future holdings received in other than machine readable form (particularly those which have significant comments in the margins handwritten and initialled by prominent people) instead of retention in paper or microfilm form. If, however, a machine readable type font was used in producing the document in the first place then the document will be capable of being both machine read word by word as information or transmitted as a picture.

A further possibility by Stage III for such forms of holdings as microfiche and video tape is the use of mini computers to locate the microfilm frame or video tape section and then move it under a reader for automatic transmission of the picture to the searcher whose access rights, of course, would have been previously checked by the central computer.

Epilogue

If the system envisaged as possible by the year 2000 for archival work comes to pass then all information held would be readily accessible – it could also be easily changed. In the Australian environment it can be confidently expected that adequate legislative safeguards against the latter form of action will always exist and so obviate any possibility of a change in the name of the Archives Office to that of the Ministry of Truth.