DOCUMENT RESUME

ED 049 781 LI 002 743

TITLE Summary of the Papers on the Integration of Computer-Based Information with Printing Techniques Presented at the Kynoch Press Seminar Held at Imperial Chemical House, London, April, 1970.
INSTITUTION Kynoch Press, Birmingham (England).
PUB DATE 70
NOTE 52p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Building Design, *Computers, Costs, Facilities, Graphic Arts, *Information Processing, Information Science, *Information Systems, Performance, *Printing, Publishing Industry, Quality Control, Standards, *Systems Analysis

ABSTRACT

Papers presented at the seminar include an introduction to information processing which describes some techniques for use of the computer; the analysis of information systems and its role in the development of information systems; the effect on systems performance of building design and environmental standards; the concept of traditional printing as a computer peripheral; and performance standards for determination of balance between cost, performance and quality to achieve technical feasibility. (AB)

AN INTRODUCTION TO INFORMATION PROCESSING

by

N.S.M. Cox, M.A., F.S.S.

Lecturer in the Computing Laboratory,
University of Newcastle upon Tyne,

and

Managing Director, Oriel Computer Services Limited,
27, Ridley Place, Newcastle upon Tyne.

Paper presented at a Seminar in London on the Integration of Computer Based Information with Printing Techniques, 17 April 1970 organised by The Kynoch Press and Oriel Computer Services Limited.

ABSTRACT

The requirements for information processing by computer have given rise to the development of techniques to meet various problems. Some of these techniques will be mentioned and described in the context of specific systems in terms of the problems of retrieval, filing and display of information records.

6

AN INTRODUCTION TO INFORMATION PROCESSING

1. Introduction

Computing traditions have often been heavily conditioned by the pressures applied to the whole field of automatic computation. The original electronic computers (and, incidentally, Babbage's 'engine') were developed for numerical computation; from this computing languages were developed to facilitate the use of these machines. It soon became clear that computers were economically viable for commercial financial administration and, what is more, purchase could be justified for the improvements provided by about 20% utilization of the computing system. This gave rise to cheap commercial computer-time in some organisations and management pressures to make use of this resource. At about the same time University computing systems were being evolved, mainly to serve the requirements for complex numerical calculation. Thus the scene was set for the development of the whole range of activities which have followed.

These developments can be grouped into a number of rather vague categories:

1.1 Language processors

Much attention has been given in the last decade to the development of 'logical' computing facilities - the development of high-level computer languages such as Cobol, Algol, Fortran, P/L1, and so on, in which the formal language in which the programmer states his processing requirements is presented to another computer program (the compiler or interpreter). The compiler than organises the computing system to carry out the required actions. A very wide range of language processors has been developed and much progress has been made in the development of formal computer languages to meet both the general and the highly specific requirements of different groups of users. Many advances have also been made in methods of compiler construction.

1.2 Applications packages

The number of types of activity for which applications packages have been developed is now almost innumerable, ranging from concordance packages handling Greek New Testament texts to on-line airline reservation systems. The main difference between a language processor and an applications package is in flexibility and generality of the language (the parameters) presented to the package to control the process and the complexity of the separate units of activity which occur. For example, a package for solving simultaneous equations may require, as control parameters (the language); only a single number giving the number of equations to be processed. The equations themselves must, of course, also be presented to the package but they constitute the information to which the process is applied rather than parameters determining the process to be carried out. In the above example the unit of process is also large - 'solve the simultaneous equations'.

1.3 Operations systems

Nearly all computers require, in addition to the hardware, a permanent resident computer program generally called the operating system, director or supervisor which, in the early stages of the evolution of computing systems, fulfilled certain minimum requirements, permitting the computing system to detect illegal activities in the users' program (such as attempting to access non-existent storage or 'rewind' a typewriter or interrogate a magnetic tape drive). In addition, monitoring facilities were included in some systems, preventing users from occupying the system for more than their allocated time, recording the accounting information associated with each process undertaken and reporting, to the computer operator, changes in system state. In the last decade, particularly recently, considerable changes have occurred in this area, due to advances in storage organisation, multi programming, multi processing and time sharing. (Multi-programming is the ability of present-day computing systems to execute processes for a number of different uses concurrently. Multi-processing is the ability to build computing systems which contain more than one processing unit. Time-sharing is the ability to produce systems at which many users may operate computer terminals, located remotely from the main computing installation, and request activity in the computing system depending on the response by the system, to the users, for the previous activity).

Present-day operating systems not only monitor processes for illegal activities and accounting, but are also capable of providing a wide range of services to the user for the management of data and intractive processing.

1.4 Storage form, organisation and management

When the main emphasis was on numerical computation, the problems of data input were trivial in comparison with those of process. The data required by the process was rigorously pre-defined - it is relatively easy to define the range of data states corresponding to a numerical value - and furthermore, for commercial financial administration with a pre-defined process, such as sales analysis, payroll processing, etc., the data states could also rigorously be defined. Access to this data, however, became more and more difficult the larger the data base (the 'size' of the total data) became. A need arose for large random-access data storage devices as the processes to be undertaken could not be conveniently formulated in terms of serial (record by record) access to the data base. As the number of times which the data had to be passed through the process section of the computing system increased, the proportion of time for which the major part of the system was standing idle also increased and there thus arose a considerable problem in attempting to attain 'system balance' - that is the relative utilization of the various parts of the system compared with the range and proportion with different process and input/output requirements which the system had to handle. In particular, although input and output peripheral devices became faster, capable of handling more characters per second, the processing capability of main processors became proportionately faster still. This, unless something was done, to alleviate this disperity, the main processor would have had to spend a large part of its 'working' life waiting for an input or output perihperals to complete their proceeding operations. One of the main methods, currently employed to overcome this difficulty, is for the operating system to intercept input and output operations to and from a user's program and arrange to store the input and output in intermediate storage so that this either can be output by the system later at its convenience or the user's program can be prevented from beginning operation until all the data for the process is available in intermediate storage. These modern computing systems require large quantities of intermediate storage both for 'spooling' (the independent input and output of data) also for random access to data. In addition, for reasons of cost, these parts of the operating system which are only occasionally used are stored in this intermediate storage and called into the main storage by the resident part of the operating system when they are required. Also it is customary to have available a wide range of applications packages, language processors and extensive library of routines (parts of programs which can be called for by the language processors) and these too are generally maintained permanently in the intermediate storage. In general, most data is obtained by the user's program by a request to the operating system from the user's program and there is therefore a need for a complex organisation and data structure so that the operating system can access any of this data as and when required. This is discussed below.

1.4.1 Storage form

The conventional forms of storage available in modern computing systems range from high-speed core storage from which a few sytes of storage can be accessed in about ½ to 2 microseconds, magnetic disc storage from which blocks of storage can be accessed in about 75 milliseconds, magnetic card file storage from which blocks of storage can be accessed in about 200 milliseconds to magnetic tape files which can only access serially at up to 80,000 characters per second - this may mean that the record at the far end of a magnetic tape might take up to 5 minutes to access. Figure 1 gives a table of aceess rates and storage costs for a range of storage media.

FIGURE 1

Storage Medium	Approximate Access Time	Approximate Cost of Storage per character		Storage Capacity (characters)
		*	**	
Fast core store	c. 0.75 micro-seconds	--	10/-	-
Slow core store	c. 6 micro-seconds	--	2/6d	-
Magnetic drum	c. 9 milli-seconds	--	6d	4×10^6
Magnetic disc (dismountable)	*** c. 75 milli-seconds	0.002d	0.1d	3×10^7
Magnetic card file	c. 200 milli-seconds	0.013d	0.5d	4×10^8
Magnetic tape	a few milli-seconds up to 5 mins.	.0002d	0.12d	2×10^7

FIGURE 1 (continued)

* excluding the cost of the peripheral device on which the data resides.

** including the cost of the peripheral device on which the data is assumed to be permanently resident.

*** making allowance for the movement of the device to locate (on average) the required block.

1.4.2 Storage organisation and data management

With large quantities of data in a complex computing system, it is generally necessary to provide, within the framework of the operating system, facilities for the storage of a variety of types of data base (the storage organisation) and a number of access facilities to this data (the data management facilities). For example it should be possible to organise records so that they can be accessed serially with data management facilities to obtain the 'next' record in main storage although the physical records may be stored together in groups (blocked) and although the system loads further blocks of records into main storage in anticipation of their being required. Similarly facilities can be provided for such serial access and, in addition, the access to records in terms of a 'name' or 'key' identifying the record. Further sophistication is clearly possible and many complex information systems use access to data through logical networks.

Developments are in hand to provide this sort of facility for large scale general data bases.

1.5 Hardware

As has been mentioned earlier, considerable improvements have been made in the speed and flexibility of computers. The 'price per unit computation' is generally decreasing. The range of types of peripheral devices has been growing, including graphical display units, graph plotters, marked sense card readers, optical character recognition equipment, computer typesetting devices and so on. Experience is growing in the ways in which such devices may be utilised within computing systems and, furthermore, with increasing utilisation their speed and reliability is also increasing.

It would seem that future advances in peripheral hardware capability are of more relevance than those advances likely in central computing systems. Perhaps the interesting features in this latter area are increased channel 'intelligence', the development of computing and data transmission networks and, in some respects, the evolution of the 'baby' computer, the central processor of which costs less than £10,000.

1.6 User requirements

Perhaps the most fruitful developments in computer-based information handling systems stem from the development of a more detailed understanding and better formalism of user requirements. Ten years ago, for example, it was thought that citation retrieval, information services (fact retrieval), library housekeeping, computer based printing and even automatic language translation were almost within the 'state-of-the-art'. The same could be said of all these fields today. However, the real development which has taken place has been, from the experience gained in a wide range of attempted applications, to expose the potential fields of application to a much more careful scrutiny, with respect to conventional practices and their justification. This has shown the need for a more careful analysis of the existing or proposed system than had previously been considered necessary, and the need for the design of much more complex systems to admit the necessary 'quirks' of such systems.

It is interesting to note that many of the more sophisticated rules - the necessary exception conditions - have been evolved because the information handling systems have almost exclusively been designed to provide service to humans, as their end product. The problems generally seem to arise in either producing results which 'seem' right to the human rather than being 'logically correct' or in producing results which will permit the human user to attain his required goal, despite his incomplete information or understanding of the product which he uses.

An outstanding example of this is the British Technology Index in which the technologist, seeking citations to relevant work, should need only to know the order in which the letters occur in the alphabet, the technical terms relevant to his own field and the ability to distinguish, both from the form and typography of the index entries, whether he is reading a cross-reference or a citation to a work.

Other examples are legion; the authority files associated with library catalogues, telephone directories, and so on: the proportion of words in a dictionary which are deliberately mis-filed: the inability, so far, to produce 'perfect' hyphenation or to translate, by computer, from English to Russian to a satisfactory level, despite the umpteen million dollars spent on research in this field.

It is at least heartening that, even though many of these problems still defy solutions adequate to the needs of some potential applications, the understanding of the technical feasibility in these areas has now become sufficiently precise for it to be unlikely that expensive mistakes will occur, provided that such experience is properly taken into account.

1.7 Methods

Many of the notable achievements in this field, some of which are outlined in the next section of this paper, have been successful due to the ruthlessness with which the objectives of the project were defined. This has been discussed elsewhere (Cox, 1969a). It is worth noting that where the methods selected were effective, they were so, in attempting to meet the limited goals of each project, by taking advantage of any well established formal principles which were relevant but in the main rigorously and pragmatically developing methods specific to the functions required by that project only. These methods developed under these circumstances are generally masked by the detail of the specific requirements of the single project. One of the main themes behind the activities of the Symplegades Research Group at the University of Newcastle upon Tyne has been to attempt to develop application-independent facilities for various components of the total 'information processing function'. This has culminated, so far, in the design of an experimental general-purpose information handling system - the 'Newcastle file handling system' - Phase 1' (Cox & Dews, 1967; Cox, 1969b). The design of a possible 'production' version of this system is in hand.

2. The Nature of Information Processing

Many of the processes for which computers have been used have required a fairly simple or a well-controlled data base - numerical values, wage rates, parts numbers and so on - however, free-form language text has rather more complex and less well defined information properties. Language text is not usually, in any meaningful sense, 'well-defined'. Attempts at formalisation of such information, such as the creation of bibliographical records, book indexes, dictionaries and so on illustrate this from the complexity of the rules for their creation, control and filing and the anomalies which still exist in such files of records despite the detailed rules.

In attempting to provide computer systems for the processing of this kind of information, it is important to recognise that not only is one concerned with a very large and often extremely complex data base but, due to the lack of formalism in the natural language parts of the data base, also one is concerned with very complex processes which must be applied to this data base to give meaningful results.

It is perhaps useful to look at some of what might be called the 'landmarks' of information processing, taken from the point of view of notable success and also from the point of view of notable failures.

A much more detailed analysis of these and other, perhaps less notable events, is given in Balmforth, Grose, and Jeffreys (1970).

2.1 Widener Shelflist Project (De Gennero, 1968)

This project has carefully limited objectives and is aimed at the publication, by sections, of the Shelflist of the Widener Library which comprised a manuscript sheaf form catalogue which was deteriorating rapidly. As the Shelflist conversion proceeds the computer tapes provide a 'data bank' of machine readable records which may be usable for other projects. Figure 2 shows a section of this published catalogue, with 'upper' case 'only' print out but using the overprinting capability of the line-printer to simulate bold faced type.

2.2 MARC Project and the British MARC Record Service (Markuson, 1965), (Coward, 1967, 1969).

These three projects (two versions of the American project and one British) have both increased the understanding of the requirements in this field and have provided an impetus for those institutions contemplating automated processing. The first American experimental service proved that 'by-product' bibliographical records were not adequate for general use as a bibliographical data base and their revised project and the British MARC Record Service have been attempting to provide a current service of centrally produced bibliographical records capable of filing, display and retrieval adequate to the needs of potential customers. Progress towards this goal, particularly in the British situation, has been better than expected and it is thought that within the next year a viable service will be available suitable to the needs of a fairly wide range of users. These projects have emphasised the need for the development of good flexible facilities for the communication of bibliographical and other information records and attention is currently being given to this point (British Standards Institution, 1970; Cox & Davies, 1970).

2.3 American National Library of Medicine - MEDLARS Project

As a result of severe congestion in the publication of Index Medicus, the National Library of Medicine decided to use computer techniques to maintain currency of publication and, because this printing requirement justified the creation of a machine readable data base, it was decided that a citation retrieval system should also be created. This data is now used by many institutions throughout the world with varying costs and quality of service. The service is rather expensive and in many cases the responsiveness of the system (the time taken to provide a list of citations) is very poor. On-line retrieval services are now being developed both at the National Library of Medicine and the University of Newcastle upon Tyne Computing Laboratory (who produced one of the most efficient of the demand search services (Barraclough 1970)).

2.4 Chemical Abstracts

This is a major series of information retrieval services based on the publication services of the American Chemical Society. The United Kingdom Chemical Information Service (U.K.C.I.S.) is now providing services in this area for the United Kingdom, but it is a little early to establish just how viable or effective these series are.

2.5 The 'Bochum' experience (Bossmeyer, 1968)

The University Library of the University of Bochum, in West Germany, was the first major European library seriously to contemplate 'total' automation. A plan was set up for total automation within five years. Now, five years later total automation, on a perhaps less ambitious scale, is expected in about a further five years.

2.6 Library Automation at Florida Atlantic University

This attempt at total automation in a relatively new library illustrates all the worst features of activity in this area - ambitious statements of intent, lively reporting of activity but without distinguishing whether these reports were intentions or realisations, followed by a lengthy period of silence punctuated by ugly rumours. In fact it turned out that the whole project had been abandoned, by direct order of the Council of the University, but the reasons for failure are still obscure, which is a pity because at least that information might have justified some of the several million dollars wasted.

2.7 The Brasenose Conference (Harrison & Laslet, 1967)

This was perhaps the most notable landmark, as far as conferences go, in the British library automation scene, for it was at this conference that about 100 librarians from both the United States of America and the United Kingdom came together and discussed their activities, aims and aspirations in this area. It was directly as a result of this conference, sponsored by the Old Dominion Foundation, that some British libraries became aware in some detail of the prospects offered by library automation and the applicability of the techniques of information processing to the activities of their own institutions.

2.8 J.C.R. Lickleider's 'Libraries of the Future' (Lickleider, 1965)

Finally, in this brief survey, it is worth mentioning the prognostications of J.C.R. Lickleider of M.I.T., the predicted 'console dialogue', 'data banks', and so on, in a book published in 1965. A point of some regret is that, although by now some of the features which he predicted are approaching technical feasibility, many librarians were misled into hoping that such facilities were 'just around the corner'. The magnitude of many of the problems, still to be faced to bring these features into realisation, is still large, despite the progress made in this field in the last five years.

3. Conclusion

In attempting to draw conclusions from this review, it is essential to look not only at the evolution and present state-of-the-art of information handling but also to relate this to the evolution and state-of-the-art of the total systems environment within which computer based information handling systems can perform. It is clear that substantial progress has been made in information handling techniques, as may be seen from the number of viable systems currently in operation. However, what is perhaps surprising is the extent to which each of these systems has been tailored to specifically the needs of the given application. Thus, each system contributes slightly to the understanding of 'how to do it' in apparently parallel situations, but hardly at all in terms of 'off the shelf' programs or systems, for other similar applications. The literature, in this area, is full of reports of 'how we did X at Y' hardly any of the reports providing to the reader information on 'what was done', 'why it was done' or 'how it was done' to a sufficient level of detail to permit the experience gained to be used in subsequent systems implementations elsewhere.

This seems to be fundamentally due to a lack of formalism in the methods adopted and a lack of formalism in the definition of the functions which it was required that the system should perform.

We have a wide range of examples of systems implementation in this field, a much fuller understanding of the general functions which need to be performed, highly evolved techniques for the design and implementation of language processors, sophisticated data management and storage systems and a crying need to integrate these into a homogeneous information processing capability. For if this can be achieved effectively, based on a satisfactory formalism of the components of the range of problem applications, then most of the current difficulties in this field will melt away. This is what we are seeking to achieve, we believe with at least limited success.

REFERENCES

Bainforth, C.K., Grose, M.W. & Jeffreys, A.E. (1969). The Computer/Library Interface: the Last Five Years - paper presented to the Newcastle Seminar on the Management of Computing Activity in Academic Libraries, University of Newcastle upon Tyne, January 1969.

Barraclough, E.D. (1970) The British MEDLARS Service, 1970 Newcastle upon Tyne, Oriel Press Ltd., (Symplegades No. 3) - in course of publication.

Bossmeyer, C. (1968) 'Katalogherstellung mit der Datenverbeitungsanlage der Universitats bibliothek' (Production of catalogues by means of the data processing equipment of the University Library) in 'Elektronische Datenverarbeitung in der Universitats bibliothek, Bochum'. Herangegeben von Gunther Pflug & Bernhard Adams, Bochum, 1968.

British Standards Institution (1970) 'Draft British Standard for A Format for the Communication of Bibliographical Information in Digital Form', 1970, London, British Standards Institution.

Coward, R.E. (1967) The United Kingdom MARC Record Service, in Cox, N.S.M. and Grose, M.W. (Eds) Organisation and Handling of Bibliographic Records by Computer, 1967, Newcastle upon Tyne, Oriel Press Ltd.

Coward, R.E. (1969). The Capabilities of the British MARC Record Service in terms of the Academic Library. Paper presented to the Newcastle Seminar on the Management of Computing Activity in Academic Libraries, University of Newcastle upon Tyne, January, 1969.

Cox, N.S M (1969 a) 'Management criteria in the design of systems for Academic Libraries'. Paper presented to the Newcastle Seminar on the Management of Computing Activity in Academic Libraries, Newcastle upon Tyne 1969.

Cox, N.S.M. (1969 b) 'The Newcastle File Handling System - Phase 1. - An Experimental information processing environment'. Paper presented to the Girton College Symposium on Computing Science, August, 1969.

Cox, N.S.M. & Dews, J.D. (1967) 'The Newcastle File Handling System' in Cox, N.S.M. & Grose, M.W. (Editors), Organisation and Handling of Bibliographical Records by Computer, Newcastle upon Tyne, Oriel Press Ltd.

Cox, N.S M & Davies, R.S. (1970) 'On the Communication of Machine-Processable Bibliographical Records.

Part 1: An analysis of the draft British Standard
Part 2: The 'Newcastle Communications Format'
Part 3: The Communications Format Access Language.

Working Papers 1970, University of Newcastle upon Tyne, Computing Laboratory.

De Gennaro, R (1968) 'Automation in the Harvard College Library' Harvard Library Bulletin, Vol. 16, No. 3, July, 1968, P.217.

Harrison, J & Laslett, P. (1967) (Editors) 'The Brasenose conference on the automation of libraries' - Proceedings of the Anglo-American Conference on the mechanisation of library services held at Oxford under the Chairmanship of Sir Frank Francis and sponsored by the Old Dominion Foundation of New York, 30th June - 3rd July 1966, London, Mansell Information/Publishing Limited.

Markuson, B.E. (1965) 'The United States Library of Congress Automation Survey' Unesco bulletin for libraries, Vol. 19, No. 1, Jan/Feb 1965.

THE ANALYSIS OF INFORMATION SYSTEMS

by

Frank Robinson

Systems Development Manager
ICI Limited, Agricultural Division

Paper presented at a Seminar in London on the Integration of Computer Based Information with Printing Techniques, 17 April 1970 organised by The Kynoch Press and Oriel Computer Services Limited

ABSTRACT

This paper will define Systems Analysis and its role in the development of systems for handling information. Fundamental analysis is necessary to achieve lasting results - there are dangers in attempting piecemeal treatments of problems.

The analyst must determine the objectives of the system and the organisational environment in which it operates. Information systems are particularly difficult in this respect that their effects are hard to trace and the uses made of information difficult to substantiate.

A new system is developed following this basic study and there are clear stages in this work. There is a need for management involvement in the process both to ensure that the proposals are sound and also to control the project.

THE ANALYSIS OF INFORMATION SYSTEMS

In the past few years the subject of information systems has had considerable emphasis in management and computer circles. It is irrelevant in the present context that the phrase has almost as many interpretations as there are adherents to it. A need for information is clearly indicated, and the uncertainty of definition probably illustrates the breadth of that need.

Such systems handle information of two distinct types; the most frequent is structured in which data items are defined precisely both in format and content. For example, a specified section of the data in a table can be defined as containing tax codes.

The other type of information, which is more properly the subject of this seminar, is in text form in which significance is indicated by context rather than by position; for example in the phrase 'for married women the most frequent tax code is 50'.

The two types of information are not necessarily unrelated. At the management level there is a great deal of textual information including extracts from structured data. Similarly tables may be presented showing information extracted from text sources.

At lower levels there will be a preponderance of structured data which may be used to prepare textual reports.

Most emphasis has been given to the presentation of structured data. Two main reasons may be given for this emphasis.

1 The ease with which significant figures can be identified giving the ability to select important facts automatically.

2 The ease with which market and production information (for example) can be presented in this way.

Another factor is, however, the difficulties so far experienced in the automatic handling of text information. This has usually meant that the information has to be prepared manually for presentation. It may well seem that the most potent text processing device so far used has been the typewriter. However the computer is beginning to have an impact on this problem and selection of relevant text is now a practical possibility with major advances in methods of presentation now being made. So far, summarising such information is not possible but this is clearly a feasible development of the selection techniques.

In the long term the techniques used in handling information are of secondary importance. It is more important to ensure that any system installed is relevant to the needs of the organisation and it is this which calls for systems analysis in order to achieve the best long term solution. The need for analysis has been increasingly apparent in all business systems in the past few years and is closely related to the increase in computer usage. Experience in computerising all types of system has clearly identified the requirement for objective study of the system before transfer to the machine. This work now tends to be done by professional systems analysts, often attached to Computer Departments. This connection should not be over-emphasised: the real justification for analysis is the clarification of the system and this is necessary as a preliminary to any change whether or not computers are to be used. Nevertheless it is true that computers form the basis of many new systems and that analysts therefore normally have computer knowledge and experience.

One effect of the analytical approach to change in systems should be to ensure that all the problems of the system are identified and fundamental solutions evolved. The importance of seeking fundamental solutions cannot be over-emphasised. Other solutions will be limited in their effect in that they may be irrelevant to the real problems or incompatible with other parts of the organisation. It is of course difficult to define possible changes in any environment over a long period and it may appear that major systems changes are automatically ineffective in the long term as they will be made obsolete by other events. Obsolescence should not occur if the design can accomodate environmental changes. The design aim should be for a system that will last for at least five years. This can only be achieved by flexibility in the various parts of the system to allow differing data and conditions to be met. It is the dynamic environment which makes the fundamental approach difficult, but this type of environment makes it more important. No other system will survive the shocks caused by the natural evolution of the organisation.

To ensure that the system is capable of surviving means that the objectives must be clearly identified. This is difficult, particularly in textual information systems but can be achieved by tracing the effects of the system throughout the organisation. These effects may consist of processing the information (e g by modifying it), absorbing for immediate use or storing for possible future use. The nature of each defined effect and the value of the information in that location has to be defined. This latter aspect is most difficult to achieve as the value of information is often more closely related to the loss due to one fact not being available than to the cost of recording each item. This is true at all levels within an organisation but increasingly so at the decision level. The analyst must make some evaluation of these intangible elements in determining the objectives of the system.

The parameters of text information systems are difficult to see. If a manual system is replaced by a computer based one the bases are entirely different. There is a different level of service for a different level of cost. To get all the benefits from such a change the philosophy of the service may have to be rethought. The most common example of this is probably in 'awareness' systems where personal services on a selective bases are much more difficult and costly to organise manually than by using a computer.

The possible extension of information sources should also be considered. The aim in all this is to identify the possible systems objectives as if existing restrictions on sources or techniques are removed. At the end of this stage of a study the analyst must have evaluated all possible uses and have identified the overall objectives.

The need for management involvement in this process is obvious. Without it the objectives of even relatively minor systems may be set in a way which is contrary to the overall organisation plan. This stage of a study is often called management analysis which indicates the need for management support and indicated the fundamental nature of the work. When the objectives of the system have been defined it is possible to initiate the projects needed to alter the existing procedures so that these objectives may be more readily achieved. It is worthwhile here to discuss the relationship between systems and projects. Most systems continue while the need for them continues; a traffic system exists whenever vehicles interact. The procedures and rules governing the system may change as the result of changes in environment (cars replace horses), objectives (improvements in safety) or technology (computer controlled road junctions). Basically, however, the system continues and projects are initiated whenever a formal study of the rules governing the system is necessary. The amended rules are often called a new system which is a useful convention but may cause confusion between the continuing system and the new procedures.

A system redesign may require one or more projects depending principally on the size of the task involved. Each project must be clearly defined both in scope and in objectives to avoid overlap between projects to enable the work to be controlled.

The work of project development falls into three main stages. These are:

a Survey

b Analysis and design

c Implementation

In computer projects implementation includes preparing and testing computer programs and this is often treated as a separate stage.

The detailed work of the systems analyst begins with the survey. The scope of the project and the systems objectives having been defined he must analyse the existing procedures in detail. This work is principally to ensure that detailed requirements (or subsidiary objectives) are defined, questioned and, if necessary, incorporated. While the initial management analysis will have classified the principal objectives there are usually subsidiary objectives covering, for instance, control procedures or local needs.

The work of the detailed analysis is carried on at two levels. During the survey the analyst checks out the feasibility of possible alternative systems. This requires that the major problems in the system are defined so that possible solutions can be explored. Those system needs which can be readily solved do not call for concentrated work at this stage.

The second level of analysis follows the completion of the survey and the acceptance of a proposal for the development of one of the possible alternatives. The detailed analysis work must now be completed and the new procedures designed. All analysis and design requires close contact with people working in the system and the ability to establish good working relationships with others is an important attribute of the successful analyst. On the user side, time must be made available for the necessary interviews and meetings on the project. The completed design is presented to the managers of the Department in a system specification. If this is accepted the new procedures are implemented. For non-computer projects this means training the staff involved, getting new forms printed and writing operating instructions. These things must also be done for computer projects; in addition the necessary programs must be specified, coded, tested and brought into an operational state.

The project development period may be short or long but always includes these stages. The end of each stage is marked by the presentation of a systems document which gives management, particularly in user departments, the opportunity to decide whether the work should continue.

These documents are -

Stage	Document
Survey	System proposal
Analysis and design	System specification
Implementation	Operating manual

In the first two of these plans for the project period will be included.

Planning is obviously most important for a large project where a number of activities must be co-ordinated. The management of the user department can obviously contribute to this operation with their experience of the problems involved in their sphere of activity and of the general management problems in planning. Techniques for estimating the time required for system development tasks are not very well advanced; it is difficult to identify the similarities between tasks in different projects. The study of document preparation in an office has no readily discernible links with an investigation into plant loadings and the range of work covered by analysts is often even wider than this. Until some basic data can be established the analyst with experience in the area being studied will be at a considerable advantage.

The development costs must also be estimated so that the project economics can be evaluated. These costs must largely be based on the times used in planning and are subject to the same possibility of error. The system operational costs are also needed for evaluation and these will be based on the designed system. The final element in evaluation is the savings or benefits expected from the changed procedures. The difficulty of placing values on the benefits derived from information handling systems has already been discussed. Some manager must attempt this no matter how difficult in order to establish priority for the work which has to be done.

As the project develops the accuracy of the various plans and estimates should improve. At the implementation date good operational figures, the result of prolonged trials, will be available on which management can take the final decision to introduce the new routines. Subsequently the operation of a new system should be checked to ensure that it is running satisfactorily both in terms of cost and user satisfaction.

To summarise this paper; analysis is a necessary part of any system change if the objectives of the system are to be defined clearly and if the new scheme is to meet them. During the development attention to detail and careful planning will help to achieve the derived results. Throughout the whole period of investigation and implementation the support of management at all levels is essential in clearing possible problems quickly and thus avoiding delay in the implementation of the new procedures.

THE EFFECT ON SYSTEMS PERFORMANCE OF BUILDING DESIGN AND ENVIRONMENTAL STANDARDS

by

Harry Faulkner Brown MC FRIBA

Principal,
Williamson, Faulkner Brown and Partners, Architects

Paper presented at a Seminar in London on the Integration of Computer Based Information with Printing Techniques, 17 April 1970 organised by The Kynoch Press and Oriel Computer Services Limited.

ABSTRACT

It is likely that the lifetime of an information system may be anything from a few months to many years. The maximum lifetime of a computer-based system does not generally exceed that of its computer, which is currently about seven years and the lifetime of a building is, perhaps, seventy years. The design or adaption of a building must therefore allow for the unpredictable changes which the development of computer techniques might impose on the environment within which these changes must take place.

THE EFFECTS ON SYSTEMS PERFORMANCE OF BUILDING DESIGN AND ENVIRONMENTAL STANDARDS

If I was asked to design a building to accommodate a system of machinery to bottle Newcastle Brown Ale, I could probably do so in the anticipation that the machinery, or something similar to it, would last 20 to 30 years. Although we are not to deal with this rather interesting subject, at once it must be obvious that drinking habits are unlikely to change a relatively fixed-function bottling plant. Moreover we are concerned in this Seminar with rapid and unforeseeable changes which might occur in systems necessitating redistribution of equipment and people within a fixed envelope of building.

A systems analyst in many ways performs the same function as an architect or any other designer. It could be maintained that the architects' method, in four stages - the statement, the investigation, the interpretation, and the recommendations, is similar to that of the systems analyst. But when the latter makes a recommendation, is not the ability of the physical space and environment a vital factor to give him or his proposed system the freedom it required? Should he be aware of the ability and flexibility of both the computer and the building hardware? In examining the total system, should he examine the interaction of the system and the cost of the space, and balance one against the other?

The most critical principle of design will probably have the greatest effect upon the use of the building in the unforeseeable future. This principle is flexibility, to which must be linked adaptability. A flexible building is the opposite to a fixed-function building. Examples of fixed-function buildings can be seen in most cities and towns of Great Britain and have been designed to meet the needs of known requirements without consideration for changing the shape, size, use or relationship of rooms. But who can predict the continuing use of a building without change? In these days of rapid development in many fields of systems or library activity, can the ratio of machines to users, or books to readers, be considered static? Can the proportion of staff space, work rooms noisy or quiet, closed stacks, and special collections beforecast with any degree of accuracy? What effect will the reconsideration of royalties, proliferation of paperbacks, microproduction, audio-visual services, and computer-orientated library system have on space requirements in offices and libraries? Will smoking, noise, background music, and television be accepted in laboratories and libraries? Will the needs of the public outweigh the somewhat despotic customs of librarianship? Will there be a move to making libraries more like supermarkets, where a book is as accessible as a packet of detergent? Will usable space for offices, laboratories and libraries lose their claim to be created as independent buildings, or will space be 'rented' in multi-function buildings? No-one can do more than guess at the outcome of any of these considerations. I suggest there is every reason to consider the design or adaption of buildings in such a way that changes are neither inhibited nor precluded, and that ability to change is made possible and easy by furniture re-arrangement rather than structural alteration.

Ideal structural flexibility can be achieved if either all columns are eliminated from the interior of the building, or if they are spaced at intervals so that they are far enough apart to give reasonable spaces for working or reading, close enough to give economy of construction, and regular enough to allow bookstack ranges, equipment, or furniture to fit precisely between them in both directions. Depending upon the size of the building plan, columns can be eliminated, with economy, from single storey buildings and from the top floor of multi-storey buildings, because superimposed roof-loading is much less than floor live-loading.

Together with this structural flexibility there is the need for flexibility of services. If a comfortable environment of well-regulated temperature, ventilation and humidity can be achieved with a constant high level of artifical illumination, and if than can cover the building overall, then redistribution of activities can be accomplished with little or no alteration to heating, ventilation or lighting services. Together these factors will give a flexible building which is easily adaptable.

Allied closely with this principle is that of extendibility. At the beginning of the planning process, projections should be made, no matter how vague, of the physical growth of the building. If the site allows for expansion, it should be considered in principle at the outset. If it is restricted and extension is impossible, the physical growth can then be directed into additional service points elsewhere. The extendible building should be designed in several stages so that it is equally acceptable at any stage. Even if the guesses are inaccurate, at worst the building will have been designed to allow for extension, and land should have been reserved in an appropriate contiguous location.

Generally speaking, an office building needs to be flexible, so that when departments or tenants change their requirements, a change of room shape and size becomes very easy and usually can be accomplished overnight by the building maintenance staff moving demountable partitions. The services of heating, ventilation, artificial lighting, and underfloor electric service runs are arranged so that little or no adaptation is required. An office building is required to carry a floor load of about 60 pounds per square foot; a computer laboratory and a library building need floors which will carry 150 pounds per square foot to allow ranges of heavy book-shelves or machines to be placed in any area. I maintain therefore that the only library building requirement which needs to differ from an office building requirement is that the library should have much more substantial floors and columns. Of course this would not apply to one storey buildings. The computer laboratory has the additional requirement of an elevated floor for flexibility and easy of access to electrical service connections.

It is interesting to compare the parallel development of office and library buildings. Twenty years ago both required windows for day-time lighting and for ventilation. For this reason, the maximum distance of a workplace from an exterior wall was about twenty feet. With two such spaces lined along a seven foot corridor the overall dimension across the width of these buildings was about fifty to fifty-five feet. To fit a substantial amount of floor space on a site required either multi-storey constructions or the wrapping of this fifty foot ribbon of floor space around courts or lightwalls.

As building technology and occupier requirements advanced, it was realised that daylighting and ventilation from windows was too variable to be accepted. The control of the artifical internal environment was developed and reached high standards. Most office and computer building occupiers accept this - librarians are beginning to accept it. The Treasury appears to accept it only if it can be created at no additional cost.

It has now been proved that it is possible to provide this high standard of environment for about £7 per square foot, provided that the building follows three principles governing the conomical air-conditioning of buildings. They are:-

1. The external surface area of the building should be as small as possible. (a cube is naturally successful).

2. The windows should be of minimum area or should be shaded so that solar penetration is minimised.

3. Open planning should be adopted to minimise the cost of air-conditioning duct work.

The first two principles need no explanation, but it is worth examining the last one. The Americans have built millions of square feet of accommodation for open-plan offices, and I am sure that most of you are aware that library design practice in the United States has adopted open planning with many beneficial results, not the least of which is the bringing together of readers and books with an almost exclusive release of the library to open access. Because of the conflicting requirements of noisy and quiet activities, in computer situations, open planning has not been favoured to any extent.

It is my considered opinion that the American open office is a ghastly, impersonal, environmental catastrophy. It offers neither oral nor visual privacy and creates an atmosphere of overall inhuman dullness. I worked in one in Toronto for twelve years. The american open libraries are somewhat better, but have quite a way to go before achieving ideal conditions. An attempt has been made to vary seating arrangements, and a high proportion of informal furniture has been introduced to make a contribution to improving library facilities.

Although for some time the Americans have led the world in office and library planning, their lead has been lost in office planning during the past three years.

I am referring to the advances in office planning which have taken place in Europe recently, which were started in Germany and given the name BUROLANDSCHAFT. I have a preference for the English interpretation 'Landscaped Offices'. This is a philosophy with a technique, of office design and layout, which claims to overcome all the traditional objections of acoustic interference that are rightly aimed at the conventional open-plan office.

The design requirements are quite simple. The plan of the office space should measure not less than seventy feet in any direction. The ceiling should not be more than ten feet above the floor, and both the ceiling and floor should have surfaces of maximum acoustic absorption. This gives a long rectangular section through the office with an acoustically inert floor and ceiling, so that hard perimeter walls and windows constitute an insignificant part of the surface of the room.

PRODUCTION AND DISPLAY STANDARDS

by

S. Albert Pace

Assistant to General Manager
The Kynoch Press, Witton
Birmingham 6

Paper presented at a Seminar in London on the Integration of Computer Based Information with Printing Techniques, 17 April 1970 organised by The Kynoch Press and Oriel Computer Services Limited.

ABSTRACT

This paper will outline the concept of traditional Printing as a computer peripheral and give comparisons of information densities between graphic arts quality character reproduction and other print-out devices. Access to characters, scanning rates and costs will also be discussed.

PRODUCTION AND DISPLAY STANDARDS

1. Introduction

Our Chairman, in his introduction, has given an outline of the events leading to the co-operation between The Kynoch Press and Oriel Computer Services Limited and the investigations we made into the use of photo-composition for our normal throughput. Our conclusions that work with a high author's correction factor was not suitable were confirmed after consultation with Mr C J Duncan but our investigations did, however, show that providing the filmsetting input was accurate there was no reason to doubt the viability of the technique.

This was emphasised by Messrs Cox and Heath in a paper given at the Newcastle upon Tyne Automated Publishing Systems Seminar in September 1969 where it was acknowledged that important differences exist in the computer handling of materials for photo-composition between the processing of collection of records and the processing of continuous text matter. It was further highlighted by the work done in the University of Newcastle upon Tyne Computing Laboratory on the general problems associated with the handling of files of fairly complex records and the provision of a general computer system for the processing of forms of records in different circumstances. The group evolved the Newcastle File Handling System which produced a photo-composition system. This composition system puts special emphasis on the handling of sequences of records and has facilities, for example, for the handling of the special classes of 'widow' which can arise in the printing of records and meets the associated problem of the bringing forward headings algorithmically as each column is ended, this heading being dependent on the position in the text at which the heading is required.

The main philosophy of this composition system is to attempt to permit the operation of a printing works to be considered as a computer peripheral operation no more frightening than output to a rather complex computer lineprinter. The approach of the commercial printing industry to the advent of computer typesetting or computer photo-composition has generally been to explore and invest in the most suitable or available system in relation to existing types of work to be undertaken. This has, in many cases, proved expensive, inflexible, frustrating and only marginally economic. We have taken a completely opposite approach and designed a system in which either one can look upon a printing works as a computer peripheral device or, alternatively, to consider the operations of a printing works to include a capability for the manipulation of the data as well as the ultimate printing.

Our approach to this field is thus, as part of the service which we offer to customers, to provide (and where appropriate stress the necessity for) a systems analysis, systems design and systems implementation of the customers' processes which give rise to the information which is to be printed. By this means we believe that it will often be possible to improve the quality and/or reduce the overall price of the product which we are providing for the customer. In addition, we intend to exploit fully the flexibility and aesthetic value of conventional printing standards.

We have already discovered that an additional benefit of such a system is a reduction in the 'turn-round' time on the production of the item and, particularly with a regular serial publication, with a commensurate increase in the currency of the material published. An example of this was the first fully computerised production of the British Technology Index, a monthly publication with an annual accumulation, in which the Editor was able to 'close-up' for publication later without any alteration in publication date giving rise to an issue almost exactly 50% larger than would otherwise be the case with all the material more 'current', very much more current than if the publication had been printed to its present typographical standards by conventional printing means. Hot metal composition would have taken two hundred hours as against the seventy-five minutes taken by photo-composition. In this case the initial processing data is undertaken by the Newcastle upon Tyne Computing Laboratory and we are only involved in the computer composition and printing.

2. Integration

The operative word in the title of this Seminar is 'Integration' and the point where the two disciplines of computer and print overlap or integrate is at the copy input stage.

Logical rules based on graphic arts techniques are written into the programs to enable the information to be presented in the most effective manner.

The computer assists the printer by making decisions on these logical rules to present the information with correct length of line, correct depth of column, etc. Graphic arts techniques cover such items as choice of typeface, i.e. Roman, Italic and Bold; size of type; indentations and general format and design. This can become very sophisticated when we include such items as hyphenation; good word spacing; elimination of widows; combination of typefaces; correct alignment; inclusion of diacritical points together with page make-up which will provide composition to the highest typographical standards. This logic of the computer programs can also assist the printer by reducing the number of keystrokes.

We accept that for certain types of printing it is not essential always to aim at these high typographical standards but despite the inadequacies of computer print out, we cannot see anyone wanting to use a typeset device simply as a computer peripheral in place of the normal line printer just because it produced a neater and more legible print out. Therefore, some graphic arts rules are necessary to make the best possible use of the equipment and decisions have to be made as to the level of quality that will be acceptable in these circumstances.

PUBLIC HOUSES : Lighting
Lighting of public houses. R M Howle *Lt Ltg.* 62 (Jul 69) p202-4 *il*
PUBLIC SERVICE VEHICLES See **VEHICLES, Public service**
PUBLIC TRANSPORT See **TRANSPORT, Public**
PUBLIC WORKS
Presidential address 1969 J Storey *J Instn Munic Engrs*, 96 (Jul 69) p183-5
PUBLIC WORKS : Engineering : Critical path analysis
Critical path analysis for an outside works division G W Booth & B Pearson *Munic Engng*, 146 (27 Jun 69) p 1287 +
PUBLIC WORKS : Engineering : Organisations
What future for municipals? *Surveyor*, 133 (6 Jun 69) p 21-2
PUGET SOUND
See
BOATS, Motor, Hydrofoil : Transport : Puget Sound
HOVERCRAFT : Transport : Puget Sound
PULP, Mechanical : Production : Bark removal : Drums : Bearings, Roller
drum barker with friction drive (Hanson Unweld by S W Hooper & Co Ltd, Montreal) *Ball Bear J* no 156 (1969) p13-15 *il*

Figure 1. An example of unjustified composition from British Technology Index

Figure 1 is a column of the British Technology Index. Note that it is set ragged at the right-hand margin and the main, if not the only reason, is that there are no word breaks which involve hyphenation, except for certain compounded words. As yet no one has written an English language hyphenation routine which is 100% perfect thus in many circumstances compromises have to be made in cases where a justified right-hand margin is required. It is interesting to note in this example that we believe that unjustified setting is preferable for two reasons:-

Firstly, in such material the natural structure of the entry is such that a very high proportion of lines end short on the right and thus, even if justified, the right-hand margin will appear ragged.

Secondly, which will be valid for many classes of index is that if long sequences of identical text succeed each other line by line, in order to achieve justification different spaces occur between the same words on subsequent lines, thus giving an impression of irregularity. (See Figure 2)

FIBRE GLASS : Reinforced plastics; Bodies : Motor cars. See MOTOR CARS : Bodies; Plastics, Reinforced : Glass fibre
FIBRE GLASS : Reinforced plastics; Hulls : Boats. See BOATS : Hulls; Plastics, Reinforced : Glass fibre
FIBRE GLASS : Reinforced plastics; Motor boats. See BOATS: Motor; Plastics, Reinforced : Glass fibre
FIBRE GLASS—NYLON 6. See NYLON 6—GLASS FIBRE
FIBRE GLASS—NYLON 66. See NYLON 66—GLASS FIBRE
FIBRE GLASS—POLYESTER. See POLYESTER—GLASS FIBRE
FIBRE GLASS—POLYESTER; Frames : Seats : Passenger rolling stock. See ROLLING STOCK, Passenger : Seats : Frames; Polyester—Glass fibre
FIBRE GLASS—POLYESTER; Housing : Capacitance detectors : Level indicators. See LEVEL INDICATORS : Capacitance detectors : Housings; Polyester—Glass fibre

Figure 2. An example of justified composition showing irregular alignment

You will also note from Figure 1 that full use has been made of three typefaces, Bold for the main entry; Roman for the brief descriptions and cross-references with Italics being used for certain parts of the information. Left-hand identations follow a logical sequence which helps the searcher readily to pick out the salient information.

I have stated that the typesetting device cannot simply be used as a computer peripheral just to produce more readable print out. I have also stated that it is possible to produce composition to the highest typographical standards which can be achieved with extra cost but we feel that compromise solutions which will present the information with excellent standards of legibility and usability is of prime importance and if this can be achieved without high computer costs then this is our objective.

Both justification and hyphenation are available and note must be taken of the considerable work which has been carried out on the hyphenation problem. This is referred to in 'Advances in Computer Typesetting' and also in the future publication of the Seminar Proceedings on Automated Publishing Systems organised by C J Duncan.

3. Input

The information can be accepted by us in three different forms:-

1. In machine generated form in either magnetic or paper tape or punched card or a combination of all three.

2. In operator punched form organised by the customer.

3. In manuscript form.

Taking the last item first, I calculate that in this category the main jobs would be those which require fairly extensive updating and require computer assistance. Year Books and Directories are examples. Even so, the number of changes must be significant to justify any proposed change from hot metal.

The second category, where the customer punches the input tapes because of editorial procedures or for other reasons, can bring Trade Union difficulties unless steps have been taken to consult with the respective parties and full agreement on principle is obtained. British Technology Index is an example of this category because indexing skill and tight editorial supervision is necessary to ensure correct coding of command signals. The Trade Union concerned, National Graphical Association, is well aware of the implications that computers can play in the future of the industry and quite rightly must protect the interests of its members. Notwithstanding this I believe the officers are sufficiently enlightened so as not to impede progress and their co-operation could be obtained with appropriate consultation.

The first category is of utmost importance, that of being able to manipulate data which is already in machine readable form or in a data bank established for other reasons. In the event of this data containing superfluous information, programs can be written to eliminate this, the remaining data being formatted for typesetting and the necessary graphic arts rules incorporated. This category very rarely requires much keyboarding and proof-reading at this stage and is therefore usually more economical than conventional typesetting.

4. Access to Characters

There are now a number of photo-typesetting devices available ranging in speed from 10,000 characters per hour to approximately 100 times that speed. An important feature of such equipment is the character accessibility varying from just over 200 to over 1000 distinct characters. To have access to 1000 sounds fine but I question whether it is necessary to have such a total as I consider that eight or nine alphabets plus the necessary 'outside' or 'pi' characters would be sufficient for any job. What is required is a system that concentrates on a wide range of capabilities but includes the ability quickly to form and include new characters or sumbols as required.

We use a Digiset as our typesetting device which normally gives access to 376 characters. These could consist of several normal alphabets, plus Greek and Cyrillic and could also include a comprehensive range of 'pi' characters which can be changed or supplemented at anytime and there is also a facility for generating special characters. The versatility of the Digiset is such that starting with four basic alphabets of capitals and lower case in normal text weight and a bold-face, these can be manipulated to produce italics and two widths each of condensed and expanded faces. Thus the four alphabets of the right typeface can be extended to forty variations.

If anyone is interested in the specification of filmsetting equipment, I can recommend Arthur Phillips' book 'Computer Perihperals and Typesetting' published by HMSO.

5. Packing Densities

I want to quote from a recent paper by E R Lannon (Ray Lannon is an Assistant Commissioner for Administration in the U S Department of Health, Education and Welfare in Washington and is recognised as an authority in computer typesetting techniques). He states that 'the printed products of computer-based information systems have been improved with respect to legibility and usability at significant reductions in total printing, binding and distribution costs......... From October 1967 to November 1969 the U S Government Printing Office composed a total of 284,090 pages using the G P O's Electronic Composing System. Using per page cost saving figures derived from an analysis of 102,000 pages, the G P O system has reduced printing and binding costs by $2,620,000. In addition, some 2,520 tons (admittedly short tons) of paper were eliminated from the distribution system. Apart from the cost savings a paragraph relates to the elimination of over 2,500 tons of paper. How can an electronic system save paper? Before October 1967 the G P O used computer print out as camera ready copy and by using their electronic system they made fuller use of every square inch of paper by packing it with more information while at the same time improving standards of 'legibility and usability'.

THE CAT SAT ON THE MAT

The cat sat on the mat

Figure 3. Composition of computer print-out and times roman upper and lower case.

If we take the words 'The cat sat on the mat' and compare the line length produced by computer print-out with the same body size in printer's type we shall find a reduction in line length of 30%. In the space taken by 73 computer print out characters, 100 same size characters could be accommodated giving more 'legibility and usability'. To prove my last statement, please look at Figure 1 on page 5.

This legibility is obtained by the use of the lower case fount with its ascenders and decenders. Your eye recognises the shape of the words and does not have to scan each letter to make it intelligible. The result is faster scanning of sentences and easier comprehension. Computer print out has ten characters to the inch and is 1/6th of an inch in depth which is equivalent to 12 point in printer's parlance. The only method of cramming more computer print out into a given area is to reduce photographically and I suggest that a 50% reduction is as far as legibility and usability will allow and this normal for a KWIC index. Reducing by 50% enables 240 characters to be contained in one square inch or approximately 40 words. In printer's type the same area would accommodate 320 characters or 54 words which means that an extra 30% can be packed into a given area.

With the reduction in size to achieve greater packing density, legibility is all the more essential and is far more significant than the 30% saving in space.

Because I have given an example based on 6 pt or 12 lines to the inch, please do not assume that I am advocating 6 pt for all setting. The size of the type used depends on many factors and these have to be fully considered. As an example, length of line is one factor which is often overlooked. It is just as ridiculous to set text of 12 pt to a width of one inch as to set lines of 6 pt to six inches. The eye does not scan easily long lines of small type and has difficulty in picking up the next line and further in the context of packing densities the length of line has a most important part to play. Take out G P O telephone directories with three columns to a page. The number of turnover lines is minimal yet if they were set narrower giving four columns the number of turnover lines would be enormously increased and space would be wasted.

6. Costs

Ray Lannon has given figures of the massive cost savings made by the U S Government which amortised a capital investment of 2.2 million dollars within two years. Can we foresee this magnitude of savings in the U K? The short answer is 'no' as there is not the same volume of information to manipulate and print but in our more modest approach we do visualise savings, savings in both money and time.

It is always difficult to assess the full value of savings derived from new projects without a full assessment of the total environment. In the context of today's proceedings I think we can fairly claim that the environment covers a range of items such as:-

1. Faster usability due to improved legibility and accessibility.

2. Lower editorial and clerical costs due to the computer manipulation of data.

3. Lower total printing and distribution costs due to the high density factor.

It is probably too difficult to assess the effect of faster usability but it must surely be significant. Increased legibility and accessibility decrease frustration and a simple example is in a telephone directory where better cross-referencing would reduce the time taken by a person in making a call. In addition internal directories amended on a loose-leaf basis normally have a very high collation cost which is completely hidden as it is dispersed throughout the organisation which the directory serves. The manual organisation of an ordered file is not only susceptible to areas of misunderstanding and accidental mishandling, it is extremely time consuming. In a similar context, how many editors have taken home at weekends masses of information for the preparation of indexes? Surely this is a hidden cost, if only on that person's leisure time.

Lower distribution costs must result from reduced weight and there is also saving in shelf space to be considered.

Against these items the costs of program writing and computer time have to be taken into account, but if the fringe benefits already mentioned are ignored, we are convinced that by applying the correct system to projects of the right size there should be a significant enough cost margin to justify its implementation. Because we are becoming increasingly aware that computers work on the 'garbage in - garbage out' principle, it follows that with computer typesetting the essential factor is clean copy. Clean input, no matter how this is obtained should give clean output, therefore, with the elimination of the correction factor, film composition costs can be lower than hot metal composition. This emphasises the fact that we cannot look at items such as print costs in isolation but must take the whole project or environment into all our considerations.

REFERENCES

1. Cox, N.S.M. & Heath, W.A. (1969)
   Integration of the publishing process with computer manipulated data.
   Paper presented to the University of Newcastle upon Tyne Computer Typesetting Research Project Seminar on Automated Publishing Systems, 7 - 13 September 1969.

2. Lannon, E.R.
   An Economic Analysis of Electronic Composition.
   Preliminary Report presented at the American University Center for Technology and Administration, Arlington, Virginia, 26 January 1970.

PRESENT DAY SYSTEM PERFORMANCE STANDARDS

by

N.S.M. Cox, M.A., F.S.S.

Lecturer in the Computing Laboratory,
University of Newcastle upon Tyne,

and

Managing Director, Oriel Computer Services Limited,
27, Ridley Place, Newcastle upon Tyne.

Paper presented at a Seminar in London on the Integration of Computer Based Information with Printing Techniques, 17 April 1970 organised by The Kynoch Press and Oriel Computer Services Limited.

ABSTRACT

Considerable variation can exist in the standard of performance of an information system. It is therefore both necessary and desirable to assess the best standard which is technically feasible and, from this, determine the balance required between cost, performance and quality. Some criteria for the determination of this balance will be discussed.

PRESENT DAY SYSTEM PERFORMANCE STANDARDS

Considerable variation can exist in the standard of performance of an information system. This variation can occur basically in three areas - the responsiveness of the system, the flexibility of the system to meet evolutionary changes and unpredictable demands for service and the quality of the end product.

1. Responsiveness

The responsiveness of a system may be measured in seconds, days or weeks. For example, an on-line information retrieval service may satisfy customers if it can respond to a user giving, say, the estimated number of citations conforming with the request profile and asking whether the user would prefer to modify the profile before the search begins, within a few seconds and produce the required citations in, say, half an hour. Similarly, a request, sent through the mail to a central agency, for processing as part of a batch, will have a responsiveness depending, amongst other things, on whether communication was by first or second class mail. The responsiveness of a publishing system may be measured in days, weeks or months and this, particularly for a 'current awareness' journal, can be of critical importance. It is quite an interesting cost-benefit exercise to attempt to assess the changes in cash value of a citation to a user as the time, since the publication of the document cited, increases.

2. Flexibility

The flexibility of a system to accommodate unpredictable demands and changes in basic requirements is also of considerable importance. A system for the production and maintenance of a printed citation file, such as a published library catalogue, should be able to meet a variety of special situations ranging from a request by the management for a statistical summary of the file based on, say, the date of publication in connection with a book stock withdrawal policy. At the same time the system should be able to meet special users' requests, for example, the now almost apocryphal example of the request for all books published in Spain in the 17th Century (Cox, Dews & Dolby, 1966).

The changes in basic requirements may require the printing of specialist bibliographies, the re-organisation of the order of the file to meet a newly detected filing situation, changes in administrative procedures such as the need for a special file to be printed out for control of accessioning, and so on.

3. Quality

The quality of a system can be assessed in terms of its capability in meeting three types of function - display, filing and retrieval. Wide variations are possible for these areas. The display may be on a computer typewriter terminal, a cathode ray tube, an upper case only line printer, a line printer with both upper and lower case characters or to a variety of printing standards. All these facilities are nowadays available and the choice is determined by the performance required by users and the costs incurred in meeting these performance standards. The standard of filing again, as can be seen from almost any computer based operation in this field, can vary from good, through acceptable, to so bad that one wonders why such a system was ever permitted to go into production.

Figure 1 shows some of the aspects of computer filing requirements. The upper part of this figure indicates seven basic options needed to meet any non-trivial 'English language' filing order. The lower part of the figure gives some of the requirements needed to meet more complex situations; for example, it may be required that Mc should interfile with Mac to prevent two separate sections in the file where in the mind of the user there is a single contiguous area. It is also quite obvious that leading non-significant words such as the definite and indefinite articles, perhaps in a variety of languages should not affect the filing. In many circumstances it is also necessary that such non-significant words and phrases should not interfere with the filing at lower levels in the filing order. There are many situations in which it is not possible uniquely to resolve the position in the file for a record and it is therefore desirable that such a special state be drawn to the attention of the editor so that he can check it and, if necessary, adjust the filing order. All these criteria and many other besides are of varying relevance in varying situations. The provision of facilities to meet these situations will depend both on the idealised requirements and also on the design and development costs of systems to provide for them as well as, of course, the additional computing cost incurred on 'run-time'.

Much has been said and written about the quality or otherwise of information retrieval services. It should be noted that there is a retrieval function in almost every system in this field. This may range from the matching of a fairly narrow personal profile in a current-awareness service and the rigorous formulation for a demand-search to the sub-division of a file into say pre- and post- 1920 material for the publication of a special catalogue. The quality of such a facility will depend both on the language available to the user for the formulation of the request and to the precision of control of the information in the data base to which the selection will be applied.

Much has been said about the merits and demerits of <u>relevance</u> and <u>recall</u> in assessing the performance of information retrieval systems. 'Relevance' is the proportion of citations returned to the user which are relevant, 'recall' is the proportion of all relevant citations which the user did, in fact, receive. Suffice it say here that, in my view, no adequate measure has yet been developed to determine the performance of an I.R. system in terms of the quality of the product. We are still, as in the margarine industry, using a 'tasting panel' ultimately to measure quality but, unlike the margarine industry, each product in this field tends to be 'tasted' independently of comparable products.

INFORMATION RECORDS

It is perhaps worth pausing here and looking at a number of information records, to obtain some perspective as to the range of materials and situations which might arise. Figure 2 shows the range of variation which might occur in a 'middle-aged' file of bibliographical records - in fact the main catalogue of Newcastle University Library. One should note the changes in style, form, content, information precision and typographical quality.

Figure 3 shows some records from a highly controlled bibliographical source - the British Museum General Catalogue of Printed Books. Here variations are, unlike in the previous examples, entirely due to differences in the documents to which these citations refer. Needless to say they pose fairly considerable problems in filing and

* WORD BY WORD or LETTER BY LETTER
* CASE INDEPENDENT, UPPER or LOWER CASE FIRST
* LETTERS BEFORE DIGITS or DIGITS BEFORE LETTERS
* FIELD BY FIELD or WITH FIELDS CONCATENATED
* VARIABLE LENGTH KEY
* IGNORING SPECIAL CHARACTERS (such as punctuation etc)
* WITH SPECIAL FILING VALUES FOR SPECIAL CHARACTERS

SOME ASPECTS OF COMPUTER FILING REQUIREMENTS (1)

Figure 1; (Upper part)

* CONDITIONAL FILING DEPENDING ON THE STATE OF OTHER FIELDS

* NON-SIGNIFICANT WORD SUPPRESSION
  (possibly for several languages)

* SPECIAL LETTER GROUP MODIFICATION
  (e.g. Mc and Mac; St. and Saint)

* NOTIFICATION OF SPECIAL STATES

* AUTOMATIC CROSS-REFERENCE GENERATION

* AUTHORITY FILE CONTROL

SOME ASPECTS OF COMPUTER FILING REQUIREMENTS (11)

Figure 1; (Lower part)

RANDALL, Walter Clark ed.

612.17

Nervous control of the heart.
Baltimore, Williams & Wilkins,
1965.

65-13484 9"

QUILLER-COUCH, SIR ARTHUR T.

L820.1 Adventures in criticism.

pp. x,222. 18 x 12 cms.

Cambridge. University Press. 1926.

Acc. No. E1,711

AMERICAN MATHEMATICAL SOCIETY

Proceedings of a Symposium in Pure
Mathematics...held in New York April 23-24,
1959.
Cosponsored by the Institute for Defense
Analyses.
Editorial Committee: A.A. Albert, Irving
Kaplansky.
Vol. 1

la.8vo. Providence, 1959.

Vol. I: Finite groups. 60-7835

AMERICAN SOCIETY FOR TESTING MATERIALS

Index to the X-ray powder data file.
Editorial staff: editor George W. Brindley...
(Special Technical Publication 48-F.)

la.8vo. Philadelphia, Pa. [1957].

[Also 2 sets of cards: Inorganic. Organic.]
Also 3 sets of cards: Inorganic & organic.
Lent to Inorganic Chem.

1957 index transferred to

56-6185-6187
57-7866-7

FARQUHARSON, ALEXANDER [Co-Author]

L710.1-BRO An Introduction to regional surveys, prepared at
the instance of the Cities Committee, Leplay House, by
Sybella Branford and Alexander Farquharson.
pp.vii,49. illus. maps. 22x14cms.
Westminster, The Leplay House Press, 1924

Acc. No. 38,948.

Figure 2. Variations in a 'middle-aged' file of bibliographical records

—— Codex Zuquinensis rescriptus Veteris Testamenti. Texte grec des manuscrits Vatican syriaque 162 et Mus. Brit. Additionnel 14.665. Édité avec introduction et notes par Eugène Tisserant. [With six plates.] pp. lxxxv 275. *Roma*, 1911. 8°. [*Studi e testi.* no. 23.]
012211.b.1/23.

—— Ἡ Παλαια Διαθηκη κατα τους Ἑβδομηκοντα. 2 vol. *Δαπανη της Βρεττανικης Βιβλικης Ἑταιρειας: ἐν Λονδινῳ*, 1914. 8°. **3015. de. 9.**
Genesis—Deuteronomy and the Psalms only. No more published.

JAVANESE.

—— De Boeken des Ouden Verbonds, in de Iavaansche taal. [illegible] [Translated by J. F. C. Gericke.] 3 dl. *'s Gravenhage*, 1854. 8°. **Jav. 94.**

Yiddish.

—— תורה נביאים וכתובים בלשון אשכנז וכו׳ [The Old Testament. translated by Jekuthiel ben Isaac Blitz.] ff. 257. 18. [*U. F. Hal-Lēvi: Amsterdam*, 1679.] fol. **1900. d. 4.**

—— תורה נביאים וכתובים מלשון הקודש נעתקים, ובלשון אשכנז נכתבים וכו׳ [The Old Testament. translated by Joseph ben Alexander Witzenhausen.] pp. 79. 150. [*E. Athias: Amsterdam*, 1687.] fol. **1900. d. 5.**

SYRIAC.

—— [*sic*] ܟܬܒܐ ܕܕܝܬܩܐ ܥܬܝܩܬܐ Vetus Testamentum Syriacè, eos tantum libros sistens, qui in Canone Hebraico habentur, ordine vero, quoad fieri potuit, apud Syros usitato dispositos. In usum Ecclesiæ Syrorum Malabarensium . . . recognovit et ad fidem codicum MSS. emendavit, edidit S. Lee, *etc.* pp. 705. [*B. & F. B. S.*:] *Londini*, 1823. 4°. **753. h. 41.**

—— [illegible]
[The Old Testament, translated into ancient and modern Syriac, the former version being the Pĕshiṭṭā, and the latter prepared from the Hebrew by J. Perkins with the assistance of A. H. Wright.]

Figure 3. Records from a highly controlled bibliographical source.

display, in addition to the special formalisation necessary to accommodate them within a data base.

These previous examples are, of course, from conventionally produced and maintained data bases. Figures 4 to 7 illustrate various computer manipulated data bases. Figure 4 (a) shows an upper-case only computer listing from the 'Union List of Periodicals in Institute of Education Libraries', and Figure 4 (b) shows an experimentally produced sample of similar material computer typeset using a 'Monophoto' photo-composer. Figure 5 shows the input data for the headings of the 'British Technology Index', but with rather inaccurate German subject words. This comes from a small trial which was conducted a couple of years ago. Figure 6 shows some citations output from the British MEDLARS System - a medical information retrieval service based on data files created by the American National Library of Medicine. Figure 7 shows a catalogue card printed on a computer line-printer with both upper and lower case characters, produced by the University of Toronto Library.

These figures illustrate some of the range of problems of information content, control and consistency which must be resolved sufficiently to meet the requirements of filing, display and retrieval of the object system.

COST ELEMENTS

A suitable balance between cost and system performance is quite difficult to establish in some areas. If one looks at a notional retrieval/S.D.I. system costs usually accrue at the following points (See Figure 8):-

Information Collection: Material for inclusion in the system has to be obtained (purchased, borrowed or at least inspected) and costs are incurred in the control of these documents.

Indexing: This is perhaps the most expensive single item, requiring scarce, highly skilled human labour in vetting the formal description of the bibliographical item and assigning subject indexing terms. This subject indexing may be restricted to the deletion of non-significant subject words from the title of the item or the modification or 'enrichment' of the title to the assignment of highly specific subject terms from a highly controlled thesaurus of terms.

Data Preparation and Editing: It is, of course, necessary to convert the information record into machine-processable form, together with attendant proof-reading and correction.

Data Base Updating and Control: The new information will have to be appropriately merged into the existing data base, usually with machine validation checks, incurring computer processing costs.

Editorial Consistency Control: All those steps, so far mentioned, will require supervision and there is generally a necessity for monitory and control procedures to be developed to maintain the consistency of use of terms, the assignment of new terms and the revision of the citations as and when it becomes necessary to increase the 'specificity' in any part of the indexing scheme.

```
EDUCACION REVISTA PARA EL MAGISTERIO
    NOS. 50-87, 1947-58, LO.

EDUCADORES: REVISTA LATINOAMERICANA DE EDUCACION
    5, NO. 30-, JULY, 1962-, CA.

EDUCAND, CONTINUED AS AUSTRALIAN JOURNAL OF HIGHER
    EDUCATION
    1-4, 1950-60, BH: HL: LO: LC: NW.
    3-4, 1957-60, RE.

EDUCATEURS
    5-26, SEPT, 1946 - APR, 1950, (LACKS: 8, 10, 17-18,
        25), LC.
    67-81, 1957 - MAY/JUNE, 1959, HL. CEASED
        PUBLICATION.X
```

Figure 4 (a). Upper-case only Computer Listing.

Title: **A.1.**
1. 1887/8. LC.
Title: **A.C.L.S. NEWSLETTER**
1-. 1949-. LO.
Title: **A.M.A.**
29-31. 1934-36; 47. no. 3-. Mar/April. 1953-. LD.
44. no. 2-. Feb. 1949-. HL.
44.no.4. May. 1949; 45.nos. 5. 7. July. Oct - Nov.1950; 46-. 1951-. (Incomplete). *NW.*
44-. no. 6-. Aug - Sept. 1949-. NO.
55-. 1960-. BR.
56-. 1961-. RE.
57-. 1962-. LO.
57. no. 4-. May. 1962-. OX.
58. no. 2-. 1963-. SW.
58. no. 6-. Sept. 1963-. DR.
59. no. 1-. 1964-. (Lacks: *59. nos. 2. 3. 4).*
KF

Figure 4 (b). Similar listing in Monophoto.

```
03/11/67   10.44 HRS.        OL1OPRE-EDIT    PRE-EDIT INPUT  BTIINPUT  / 01

        NAHRUNGSMITTEL!*LT*ZUBEREITUNG!.QUALITAT**LP*KONTROLLE!.STATISTIK*EXP*.STA
0000002 FREQUENZ!*MULTIPLIKATOREN!.WECHSELRICHTER!.SCHALTER.*LP*EINPOLIGE*EN*
0000003 GALLIUM**LP*ARSENID!.RAUM**LP*LADUNG**LP*WELLEN*RP*.ELEKTRON*EXP*.ELEKTRON
0000004 GAS..STADT!*LT*HER*STELLUNG!.KOHLENWASSERSTOFFE!.UMFORMUNG!.KATALYSATOREN*E
0000005 GAS*.CHROMATOGRAPHIE!*LP*FRUCHE!*LP*ANSAMMLUNG!.MOLEKULARE*.SIEBE*EN*
000000A GENERATOREN..ELEKTRISCHE..D.C.!.KLAMMERN!.STAHL!.PRESFORMEREI!.ZUFUHRUNG.*
......  ATISCHE*E**
0000007 GERMANIUM!.KRISTALLE.*LP*EINZELNE!*LT*HERSTELLUNG!.ZIEHEN!*LP*UBERZUGMITTE
0000008 GOLD!.BERGBAU!.AUSBILDUNG*EN**.TECHNISCHE AUSBILDUNG!.MANAGEMENT!.PLAN.SPI
......
```

Figure 5. Input data for British Technology Index headings in German.

```
GREINER AC NICOLSON GA
  NEW SIDE EFFECTS IN PROLONGED CHLORPROMAZINE THERAPY.
  ENG

  CANAD PSYCHIAT ASS J 10 109-11 APR 65
   *CHLORPROMAZINE TOXICOLOG   DRUG THERAPY   MELANIN  *MELANOSIS   METABOLISM   PATHOL
   *PIGMENTATION DISORDERS   PSYCHOSES   TOXICOLOGIC REPORT   VISION


HAUFTOVA O SLAVIICEK J
  CERULOPLASMIN IN SERUM IN LIVER DISEASES
  CZ
  CERULOPLAZM'N V SERU U JATERNICH CHOROB.
  CESK GASTROENT VYZ 20 145-53 APR 66
    CERULOPLASMIN  META   DIAGNOSIS, DIFFERENTIAL   FEMALE  *HEPATOLENTICULAR DEGENER
   DIAG   MALE


HAUFTOVA O SEIDLOVA V SLAVIICEK J HINARIKOVA E VALACH V
  WILSONS DISEASE WITHOUT NEUROLOGICAL SYMPTOMS
  CZ
  WILSONOVA CHOROBA BEZ NEUROLOGICKYCH P(RIZNAK)U.
  VNITRNI LEK 11 105-12 FEB 65
    ADOLESCENCE   DIAGNOSIS, DIFFERENTIAL  *HEPATOLENTICULAR DEGENER  *LIVER CIRRHOSIS


HEINRICH F KALTENBACH M
  WILSONS DISEASE IN THE FORM OF A PANHEMOCYTOPENIA
  GER
  MORBUS WILSON UNTER DEM BILD EINER PANHAMOZYTOPENIE.
  MED KLIN 60 499-502 26 MAR 65
    BLOOD   BLOOD COAGULATION TESTS  *COPPER   DRUG THERAPY   HEMORRHAGIC DIATHESIS  *
  * PENICILLAMINE  *THROMBOPENIA
```

Figure 6. Citation output from British MEDLARS system.

```
                Architecture of the ancient
NA                civilizations in color
210           Cichy, Eodo, 1924-
C513              Architecture of the ancient
1966a         civilizations in color; Mesopotamia,
              Egypt, the Indus Valley, the
For copy      Megaliths, the Hittites, the
and [illegible].  Minoans, the Mycenaeans, the
see [illegible]in  Etruscans, Central and South
entry         America. [Translated from the German
              by A. K. Bakker]  [London] Thames &
              Hudson [1966]
                                        see next card
                    1  ul                    66022245
```

Figure 7. Catalogue card using upper and lower case characters from a computer line-printer.

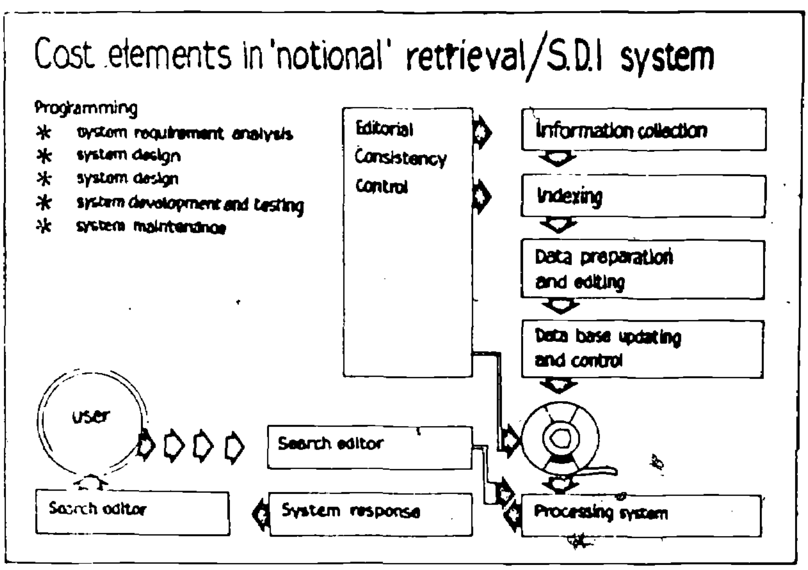


Figure 8. A Notional Retrieval/S.D.I System

Search Editor: In most schemes, it is generally desirable to provide a 'human' interface with the user - a 'search editor' whose function is to interrogate the user and re-formulate the user's, often quite hazy, request in terms of the specific measuring assigned to subject terms by the indexes.

Processing System: The processing system and associated administrative computer programs have, of course, to be designed, developed, tested and, often forgotten, maintained. Costs accrue here in both personnel and resources used.

It is fairly clear that the performance of the system will depend critically on the level of investment in these activities. The service to the user will be affected by the precision of indexing and editorial control of the data base and its input; the rate of emendation and currency of the data base, the sophistication of the processing system and the amount of support available to the user in formulating his request. The design or selection of a system to meet the requirements of a situation on the most cost-effective basis is, to say the least, non-trivial!

Display standards are perhaps more amenable to costing. It is not too difficult to provide costs for any proposed standard of display; however the benefits which might accrue from any given standard of display are more difficult to judge. The paper by Pace (1970) has given us some interesting technical detail on how costs can be minimised in some situations. It may perhaps be interesting to look at some of the different ways in which a small S.D.I. (Selective Dissemination of Information) bibliography can be accumulated and printed to meet the varying requirements of different kinds of user.

Figure 9 shows a scheme in which successive accumulation into 'quarters' and successive quarterly accumulation attempts to minimise the number of documents to which reference needs to be made by the user and assumes that the user is not specifically interested in 'this month's' references - rather he is interested in all references so far. The costs shown in Figures 9 to 13 are based on a computer-based publishing system with the information already in machine-readable form. In these schemes publications have been costed with binding appropriate to the size of each issue and in this example a '1 line per entry' index has been included with each of the quarterly accumulations.

Figure 10 shows an alternative scheme in which only the current month's references are given in each issue, but the one line index is cumulative, thus giving a monthly current awareness service with reference access to previous issues within the year. Note here that the scheme presents, as in the annual volume both the total cumulative entries for the year and, separately, the December entries.

A further scheme is given in Figure 11 where, similar to the previous example, only the current month's references are given, but in this scheme the cumulative index only appears quarterly.

A fourth alternative is given in Figure 12 which is an extension of the previous scheme, here presenting quarterly a cumulation of both entries and index.

'NOTIONAL' PUBLISHED S.D.I. SERVICE

SCHEME 1 (150 entries per month)

1	JAN	1
2	FEB	1 + 2
3	MAR	1 + 2 + 3 + Index (1 - 3)
4	APR	4
5	MAY	4 + 5
6	JUN	1 + 2 + 3 + 4 + 5 + 6 + Index (1 - 6)
7	JUL	7
8	AUG	7 + 8
9	SEP	1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + Index (1 - 9)
10	OCT	10
11	NOV	10 + 11
12	DEC	1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + Index (1 - 12)

Annual Cost

£2750

Figure 9

'NOTIONAL' PUBLISHED S.D.I. SERVICE

SCHEME 2 (150 entries per month)

1	JAN	1
2	FEB	2 + Index (1 - 2)
3	MAR	3 + Index (1 - 3)
4	APR	4 + Index (1 - 4)
5	MAY	5 + Index (1 - 5)
6	JUN	6 + Index (1 - 6)
7	JUL	7 + Index (1 - 7)
8	AUG	8 + Index (1 - 8)
9	SEP	9 + Index (1 - 9)
10	OCT	10 + Index (1 - 10)
11	NOV	11 + Index (1 - 11)
12	DEC	12 + 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + Index (1 - 12)

Annual Cost

£2285

Figure 10

'NOTIONAL' PUBLISHED S.D.I. SERVICE

SCHEME 3 (150 entries per month)

1	JAN	1
2	FEB	2
3	MAR	3 + Index (1 - 3)
4	APR	4
5	MAY	5
6	JUN	6 + Index (1 - 6)
7	JUL	7
8	AUG	8
9	SEP	9 + Index (1 - 9)
10	OCT	10
11	NOV	11
12	DEC	12 + 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + Index (1 - 12)

Annual Cost

£1865

Figure 11

'NOTIONAL' PUBLISHED S.D.I. SERVICE

SCHEME 4 (150 entries per month)

1	JAN	1
2	FEB	2
3	MAR	1 + 2 + 3 + Index (1 - 3)
4	APR	4
5	MAY	5
6	JUN	1 + 2 + 3 + 4 + 5 + 6 + Index (1 - 6)
7	JUL	7
8	AUG	8
9	SEP	1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + Index (1 - 9)
10	OCT	10
11	NOV	11
12	DEC	1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + Index (1 - 12)

Annual Cost

£2585

Figure 12

NATIONAL PUBLISHED S.D.I. SERVICE

(150 entries per month)

with successive annual accumulations

	Scheme 1	Scheme 2	Scheme 3	Scheme 4
YEAR 1	£2750	£2285	£1865	£2585
YEAR 2	£3450	£2990	£2570	£3290
YEAR 3	£4160	£3695	£3275	£3995
YEAR 4	£4860	£4400	£3980	£4700
YEAR 5	£5560	£5105	£4585	£5404
YEAR 6	£6270	£5810	£5390	£6110

Figure 13

Figure 13 summarises the costs for these four schemes and shows how the costs increase if further cumulations are required. These costs provide for cumulation of both the entries and the index at the end of each year.

From these approximate figures, one can thus compare the costs incurred in some of the choices of user service which could be offered.

Present Day Performance Standards: A number of substantial technical limitations exist in this field. There are also a number of formal limitations; however, it seems that in many respects the performance standards attainable from computer-based systems are technically superior to conventional systems in either 'turn-around-time' or quality. As was mentioned earlier (Cox, 1970), our philosophy has been to develop a 'high-level' integrated system embodying general purpose facilities to the best technical standard which can be attained. In addition, of course, it is necessary to arrange for this system to link appropriately with other systems and computer languages, permitting the inclusion of special facilities designed specifically to meet the requirements of a particular project.

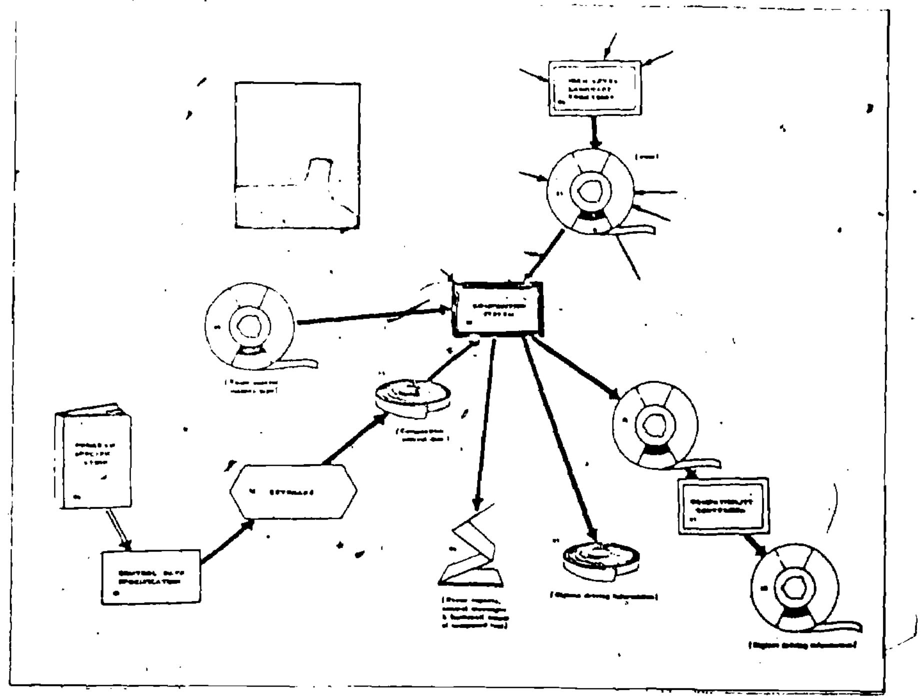
The way in which such a system works can be seen in Figure 14, which shows some of the ways in which packages can be linked to a composing and printing system. In this example, a number of routes through the system are possible, permitting revision and editing of data, and manipulative access to the data bases by high level language facilities to create the master file for composition and further high-level systems to modify this file before passing the data to the composition system which is, in a sense, also a high-level language capability.

Other, simpler, routes can be developed to meet more straightforward situations and in the background is the general capability of the various components of the total system.

One advantage of such an approach, not mentioned so far, is the convenience of these logical facilities for communication, both with users and amongst members of the programming teams.

What then are the best standards technically attainable in this field? A fairly good filing standard is attainable in many situations; however there are circumstances where the discrepancy between a logical order and the order which the reader expects to see is noticeable. Perhaps the most serious worries about filing orders is that there is no good measure of quality of filing, and that it is always possible to 'lose' a section of a file because of slight mis-filing. In keyword-in-context situations, for example, algorithmic 'tail-cropping' (the removal of the tail of a word to file only on the root) is still not entirely satisfactory. It should, perhaps, be remembered here that most information files are subject to continuous growth and that the complexity of the filing algorithms applied must increase with increasing file size, to maintain the same average size of sub-division within the file.

The standard of performance in display, depends both on the sophistication of the computer algorithms and, more particularly, on the capability of the display peripheral system. Typewriter terminals and C.R.T. display terminals are



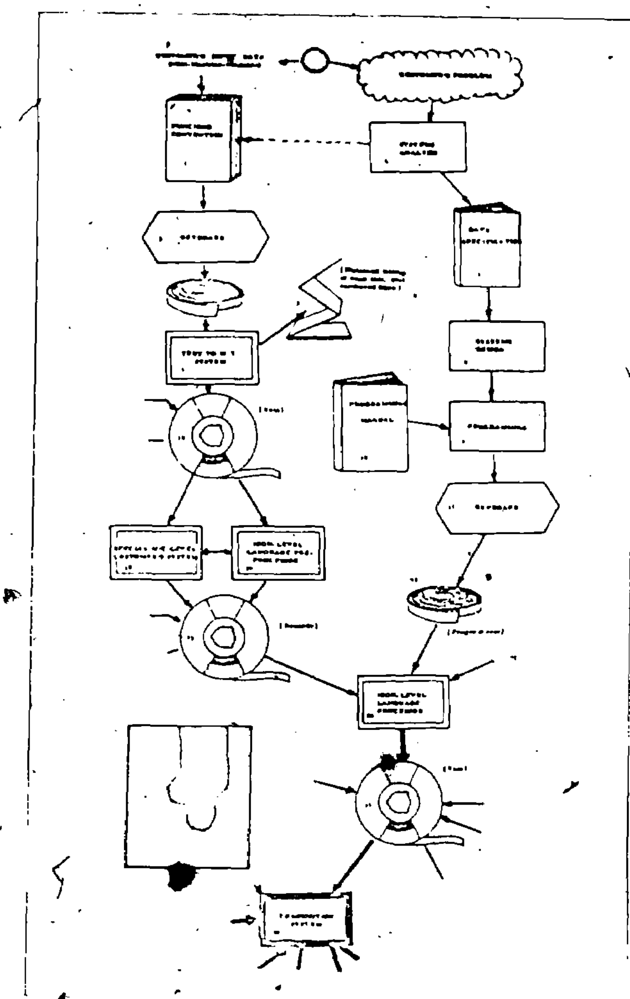


Figure 14. Illustrating some of the features of an Integrated Publishing and Composing System.

available and used; however, for some uses they have a rather restricted character set. New character display terminals are becoming available with many more distinct characters though one cannot generally with these introduce characters of one's own choosing. It is interesting to note that, for some purposes, these terminals appear to be much less tiring to use than the equivalent typescript. A further limitation is the maximum number of characters which can be displayed at any one time. Many computer line-printers now have fairly extended character sets, though the majority in use at the moment are upper case only. Usually, large character set printers are much slower than those with the restricted character sets and this does mean increased processing costs.

A wide range of standards are attainable in respect of printed material. Costs too can vary enormously. However, on the whole, and with the exception of such problems of hyphenation quality, a situation now exists in which the only ways in which the user should be able to detect the involvement of the computer in the manufacture of the printed article are the high degree of consistency, the speed with which the item can be produced or the unusually wide range of founts, point sizes, alphabets, and special characters which can occur. One is now in the position where, in such circumstances, the printing capability is much in excess of the equivalent conventional standards.

Retrieval standards are again more difficult to assess; however there are a number of parallels, such as U.K.C.I.S., MEDLARS and a number of proprietary management information systems.

It should be noted that all these standards and criteria are subject to the quality of the information in the data base and the 'linguistic' algorithms which can be applied to this base.

CONCLUSION

It is now technically feasible to produce a wide range of systems in this field, performing often to a better standard than is attainable by conventional methods. This is particularly true for some classes of printed records.

Considerable choice is available to the user in the responsiveness, flexibility and quality of product of the system to be adopted; however cost-benefit exercises are generally difficult.

An estimated cost for the printing of 500 copies of the Newcastle University Library catalogue (which contains about 300,000 records) from an existing machine-readable data base is in the region of £18,000. A librarian costs something in the region of £3,000 per annum to maintain. What does the University prefer - a copy of the catalogue on everybody's desk or a librarian for six years? This is the sort of problem which must be faced in this field.

REFERENCES

Cox, N.S.M., Dews, J.D. & Dolby, J.L. (1966)

The Computer and the Library - the role of the computer in the organisation and handling of information in libraries.

1966, Newcastle upon Tyne, University Library (Library Publication No 4)

Cox, N.S.M. (1970)

An Introduction to Information Processing. Paper presented to the one-day Seminar on the 'Integration of Computer Based Information with Printing Techniques' 17 April 1970.

Pace, S.A. (1970)

Production and Display Standards. Paper presented to the one-day Seminar on the 'Integration of Computer Based Information with Printing Techniques' 17 April 1970.