THE MOVING FINGER,
OR
THE FUTURE OF INDEXING*

R. D. GEE

The theme of the Rubaiyat is the mystery of existence, doubt of the future, and belief in the present. I do not accept this philosophy. We cannot alter the past, but we can record it; and within reason we can use these records to anticipate and direct the future. We are irresponsible if we refuse to move from the interim present. Besides using the Rubaiyat to introduce comment on futurology, I will paraphrase it to provide a pertinence to indexing:—

'The moving finger hunts; and having hit,
Flips back: nor all thy piety and wit
Shall lure it back to check but half a line,
If what was indexed proves a proper fit.'

In marshalling my thoughts for this paper I met certain fundamental problems. In analysing them I found that I had created the very framework of the presentation. For instance, was I to state a future for indexers as I know you today—or for the applications of the craft of indexing, whoever is responsible for them? The overriding enigma is not conviction that indexes will exist, but the physical format of the material requiring indexing. This leads to questioning the aptitude of present methods for media yet to be developed.

I mention this because in the past five years I have worked with both a computer company and with a publishing group. My present work is related to research on printing and communications technology of future potential for publishing, and I am no longer prepared to accept paper as an a priori commodity for information transmission. Once one doubts the importance of crushed trees as the raw fuel, one begins to realise that all prognostication relies on the same confidence as that which guides an arm down a rabbit hole—to be bitten by an owl, which has no logical right to be there.

The future will be full of things that have no logical right to be where they appear—by today's concepts, that is. After all, we now accept that Stonehenge was built as an astronomical index, interwoven into a totally different culture to ours. To the builders of Stonehenge, Salisbury Cathedral has no right to exist.

A conference has just finished in New York, devoted to computers in printing, called COMPRINT 90. One paper was entitled 'The future isn't what it used to be' from a science fiction short story by Arthur C. Clarke®. The author, Arthur Gardner, introduced his paper with one of Clarke's predictions:

'Within ten years many households will be equipped with communications consoles. The con-
sole will consist of a television camera, a t.v. screen, a computer terminal, microfilm facilities, and a newspaper facsimile receiver—all linked to computers which, in turn, will be linked to communications satellites... Any newspaper in the world, from the Wall Street Journal to Pravda, would be available to any user of the console, no matter where he lives. And, thanks to microfilm storage and retrieval, any newspaper ever printed (back to the 18th century) would also be available, as well as any book, magazine or pamphlet worth saving. Such material would be called up from a global electronic library—a central data bank linked to the total electronic nervous system of the world.'

According to Arthur Gardner, Clarke's vision is as ridiculous as sending a man to the moon. The printing world has a future—but that future isn't what it used to be.

Certain technologies have remained stable because their physical presences have dominated an environment that in turn has conditioned them. Printing, as the oldest technology, has not changed much since the first steam presses of The Times—installed between the Retreat from Moscow and the Battle of Waterloo. This may have been because of the dominance of mechanical engineering, with electrical engineering as a support technology. The last decade, under the Svengali of space travel, has dictated that physical energy be transmitted with as few moving parts as possible. We are now at the beginning of not just an electronic age—but an opto-electronic age.

Don't accept my views as a prelude to a Wellsian technological Utopia. Our descendants will expend colossal effort and resources in converting the planet Venus into a human ecology, but our own deserts will still be deserts; our seas will still be inviolate beneath the surface; and the need to find a fresh planet will be stimulated by the polluted ruin that will be Earth.

In the meantime our future will be governed, not by the mere transmission of power, but of energetic senses. Sound has been conquered; vision has some problems left; and the future will tackle touch, smell, and intelligence. When we have identified the elusive sixth sense, that will be transmitted—probably into the fourth dimension; and with suitable trumpeting from the fourth estate. All these labels refer to channels of information. The progress of man is measurable by his ability to transmit information, to record the process, and to recall it at will. Therefore, before I settle down to more mundane matters, I feel it necessary to break your conditioned responses to the printed word as it represents a concept. In fact, I read recently that the retreat from the word, or Gutenberg rut, is gaining momentum. Just consider that if a smell can be transmitted, stored, and retrieved—what logical processes can be used to index it? Somehow the brain manages it, and one day the method will be known and applied.

Since I am expected to foresee the passage of the moving finger, I intend to preface with some opinion on the subject known as 'technological forecasting'. This can be facetiously described as using statistical methods and computers to analyse chicken bones. One well-publicised system, developed by the Rand Corporation, is the Delphi technique—defined in one of its reports as 'a methodology used for the elicitation of opinions of experts'. Isolated specialists forecast dates by which stated achievements will have been made. The correlated results are graphically presented as tabular catacombs of sarcophagi, the peak of each coffin representing the consensus of expert opinion. I choose this metaphor to remind me not to be misled by the cult of the expert—who too often listens to his own kind in order to cast more bones. I try to spot the living flesh. Future, like truth, is not shaped to fit into the hand of any person—or even into the cupped hands of a collective esotery. The problem is not new, and I remind you of a Shakespearean couplet:

'The mortal moon hath her eclipse endure
And the sad augurs mock their own presage.'

The Archivist of the United States recently prescribed the possible future of hard copy. I define this as 'information presented through conventional symbols on a paper-like base'. I use his comments here because
of their relevance to any discussion of the future.

In 1950 (when people spoke in awe of electronic brains) 95 per cent of our records were on some form of paper. By the year 2000 this ratio will not be more than 50 per cent. Apart from the already known photographic film, we will have film bases which are diazo, vesicular, photochromic, ultra-high resolution, dry silver, or photopolymer. Paper will have been replaced or supplemented by thermoplastics or other plastics, either metallic- or magnetic-oxide based—which may have been imaged by video, hologram, or by magnetic-optical techniques.

To this, not only must be added materials and techniques yet undiscovered—but storage media intended purely for machine processing, such as tapes, discs and cassettes—and a yet undefined range of information to which I can only allude as ‘soft-copy’ (that is, disposable and short-lived) or ‘palimpsest’ (instantaneous display without copy).

It should be apparent that indexing, whether computer-assisted or not, has yet to face its greatest problems; and perhaps its greatest exponents are at present sucking tiny index fingers. Before you contemplate these problems, however, consider the following statements from experts:

‘Space travel is utter bilge.’—Dr. Wooley, Astronomer Royal, 1956.

‘As far as sinking a ship with a bomb is concerned, you just can’t do it.’—U.S. Rear Admiral Woodward, 1939.

‘Talking on the screen destroys the illusion. Devices for projecting the film actor’s speech can be perfected, but the idea is not practical.’—Thomas Edison, 1926.

In every case, each expert was proved dramatically wrong within two years, although curiously these very events had already been rehearsed in some practical form. Edison himself had demonstrated the first public sound and motion experiment with his ‘Kinetoscope’ in 1894; while in the very year of this quotation the John Barrymore film ‘Don Juan’ was released with an introductory speech from the screen and a synchronised musical accompaniment throughout.

And a word for the others. In 1877 the Tay Bridge was opened, designed badly, incorporating false data submitted to Sir Thomas Bouche by, among others, the Astronomer Royal of the day. (There must be something about Astronomers Royal.) As self-qualified eulogist and Dundee bard, William McGonagall burst forth:

‘Beautiful Railway Bridge of the Silvery Tay! I hope that God will protect all passengers By night and by day, And that no accident will befall them while crossing The Bridge of the Silvery Tay, For that would be most awful to be seen Near by Dundee and the Magdalen Green.’

Two years later it fell down in a storm, together with a train and passengers, and was witnessed by many from ‘nearby Dundee and the Magdalen Green’.

Still, back to indexing.

The immediate, but improvable, present can be represented by those periodical indexes for which I have no responsibility whatever, because I am no part of the operational side of publishing. I am more concerned with the product as a consumer. With weeklies like New Scientist, European Chemical News, Computer Weekly, Electronics Weekly, or even Practical Gardener, Woman’s Own, or Motor Cycle I feel that the Lynch programme could have some contribution to make. Despite the joint attitudes of editors and circulation managers I believe that an index can benefit circulation, and ultimately the advertising income. If regular readers were provided with an index, they may view their weekly expenditure as a con-
tribution to a usable reference set—which, of course, must have no issues missing.

Unfortunately, editors see no value in yesterday’s material, since easy retrieval may prevent a repetition in the future of a fresh feature. In the consumer magazine field, too much is made of the value of point-of-sale display, and not enough of the dynamic encouragement of regular subscription. The overlying economic principle for this attitude is historically justifiable, for as the domestic purse strings are tightened, the newsagents’ orders are curtailed. It is up to publishers to ensure that the first cuts will be made elsewhere, because every system, however ideal, always faces fluctuations of economic stress. If I had the opportunity and energy to argue for an indexing policy within my company, I would pin my argument along those lines—not on utilitarian ideals. When the publisher of a serial refuses to recognise the value of an index, he has publicly admitted that his product has no value. Therefore his public cannot be blamed for purchasing from week to week, and breaking the habit on the shortcomings of one issue.

Certain weeklies could benefit by quarterly, bi-ennial, annual, and possibly further cumulated indexes, since services like British Technology Index and British Humanities Index cannot cope. In many cases the quantity of text in news items is too small for high-level indexing, although the information quotient is high enough for original editorial selection. Publishers see themselves as harvesters of ripe corn, rather than as purveyors of information. They should repeal their Corn Laws.

With monthlies I have less belief in the need for indexes, although I would like to use more five-year cumulations. Monthlies are well covered by commercial and institutional abstracts and indexes, by services which are professionally constructed, which are subject to less delay, and which cover other relevant publications. Where indexes to monthlies exist as separates, they are not amenable to sophisticated techniques. For example, when twelve issues produce a total of eighty or so articles, an annual index has value only for library binding or to support microfilm files. If journals are kept boxed, which I prefer, a separate index is useful as a guide to missing copies, although even then an annual contents list is better. In other words, periodicals require individual indexes when the annual quantity of information is high, and such indexing is now possible by the use of data processing. The use of computers does not presuppose the by-passing of intellectual labour. The Lynch programme, for instance, demands a high level of analysis and heavy use of key-punching equipment.

The fault of individual journal indexes, for a librarian, is that one supposes the enquirer to have identified a source for his problem first. This is all well and good if you expect to find it there, because the journal is so specialist, or because it was remembered from an earlier reading. This sort of information work is too hit-and-miss, and publishers do not help by automatically providing indexes as a natural part of the subscription—even when they do exist.

You may reasonably wonder why I stand here and criticise my own industry, stating what is needed, but not converting from within. And since you are reasonable people, you will require a sort of reason. Publishing is a curious activity. Although tied to the first form of mass production, it stands on a cottage industry concept. Authors are often not employees, and are responsible for equipping themselves with the information they need—which probably accounts for their pressures for a Public Lending Right. Editors, illustrators, and indexers are hired in short doses like consultants or building labourers. A publisher does not maintain a reference library for staff, and the closest sort of information file is an inbred and empirically handled accumulation of cuttings previously generated by that publisher. Outside of that, ‘stringers’ are used—or the burden thrown at public libraries. And this accounts for some librarians’ reaction to a
Public Lending Right. In general, all information is bought, begged, or scrounged from outside, and once used, is forgotten—so that the process can be repeated in the future. Although the situation is improving, you can appreciate the difficulty of spreading the gospel in such an environment. Arts-graduate-bound book publishers, failed-graduate-riddled magazine-houses, and anti-graduate newspapers all share one thing in common. They all survive by selling information, but they have little comprehension of it as a commodity, and none at all of it as an asset. The only consolation that I can masochistically offer, is that if the lesson is not learned within a decade they will not survive, and the commodity will be more efficiently handled by a new generation of alert enterprises—and electronics has nothing to do with it.

As a user of indexes, I have mixed feelings on the universal compendia like British Technology Index or Engineering Index. Headings become so complex that they face changes from year to year. Thus time-expansive searches become maddening. Better terms are discovered as one progresses, so constant re-searching becomes necessary. In highly specialist fields they become cumbersome, despite excellent design philosophy. Consider a typical pattern: assuming a search for a specific application of a programming language, FORTRAN IV, the searcher comes to grips with DIGITAL COMPUTERS: PROGRAMMING: LANGUAGES: FORTRAN: IV. This becomes more complicated if he seeks a sub-set when the original indexer is unaware of the hierarchical structure of some languages. For example, if COMPTRAN is an application sub-set of FORTRAN, is COMPTRAN IV a sub-set of FORTRAN IV, or a development of COMPTRAN I to III? This is where the subject abstracts win the day. It also indicates the need for the indexer to have a good grasp of the text.

It is regrettable that those publishers providing annual indexes to particular periodical titles have not appreciated the value of cumulation, nor the need to concentrate on the easier compilations for companies, products and trademarks from those infuriating pages at the back of commercial technical journals. The only real cumulation I know appearing regularly is that to the Economist Intelligence Unit's Retail Business. Each annual index covers back to issue 1, with 1-148 having just been published. Unfortunately the construction is crude, being title-based rather than subject; it is without cross-references, and it includes reports that have been superseded—without so indicating.

Certain periodicals have produced bumper bundles, such as Scientific American and Engineering, but these are rare and not part of a regular policy. Computing Reviews produces fairly regular cumulations for two or three year clumps—but true to the new traditions, they are in KWIC format, which does not endear me to them.

I also have to live with abstracts in printing, like that of PIRA whose latest annual index covers 1967—while they use their computer to produce an indifferent monthly printout of articles relevant to newspaper production. Since technology does not progress in annual cycles, I see no reason why subject indexes to abstracts should not cumulate in at least three-year compilations. On the other hand, an American equivalent of PIRA, the Graphic Arts Technical Foundation, does not index its monthly abstracts at all—and consequently wastes a great deal of work, because no one in their right mind would conduct a retrospective search through their chunky product.

Since I take over fifty subject indexes and abstracts in all fields of our research, I attach little value to the individual journal index, beyond that of a useful ‘fill-in’ to literature searches. But consider the value of a ten-year cumulative index to the Library Association Record or The Indexer.

The reason for all these shortcomings is predominantly a shortage of full-time professional indexers.
The near future

I have mentioned some simple realisable improvements to today’s efforts, which may need data processing assistance to handle the volume in sufficient time to make the progress worth-while.

There are now several comprehensive indexing services for news material that did not exist five years ago because no commercial need was recognised. Within five years these may be computer-driven. Research Index, with its crude production, disorganised subject headings, and lack of cumulative indexing, would be an incredible tool with computer support. It should compare itself with its American counterpart of Funk & Scott.

If any news index deserves more sophisticated techniques to match its impressive appearance, it is the Official Index to the Times—although it has improved enormously in the past decade. Its most unforgivable fault is its tardiness. Too much editorial effort is thrown away by a production insistence on quality printing and prestige binding, thus destroying the prime motive of a news index—namely topicality. There is no reason why a paper-covered edition could not be produced, and cut out the eight-month delay which customers at present tolerate for relevant indexes. At the moment I receive my microfilm Times six months before they are indexed; this seriously affects the value of microfilm storage. I would suggest that the Times takes a close look at the H. W. Wilson publishing policy, and they may find their microfilm sales improving.

I will return later to what may be called ‘case-hardened’ book-publishing stupidity, but to introduce some coverage of computers I move on to a quotation from Computing Reviews, vol. 2, No. 9, September, 1970:

‘In striving for a speedy method of organizing an index, the question arises as to which of various possible schemes is adaptable to fully automatic processing. Clearly, some means of ordering is required that is based on criteria extracted from the text itself rather than assigned in accordance with human judgment. The simplest format of a quickly assembled index might be an alphabetizing listing of keywords, very much as in the index of a book. In dealing with a variety of subjects, the significance of such single keywords could, in most instances, be determined only by referring to the statement from which the keyword has been chosen. This somewhat tedious procedure may be alleviated to a significant degree by listing selected keywords together with surrounding words that act as modifiers pointing out the more specific sense in which a keyword has been applied. The added degree of information conveyed by such keyword-in-context indexes, or ‘KWIC Indexes’ for short, can readily be provided by automatic processing. Keyword-in-context indexing may be carried out on various levels, depending on the purpose an index is to serve. The process may be applied to the title of an article, its abstract or its entire text.’

H. P. Luhn.


Before I move into the environment in which the computer may become a natural part of indexing, I will mention some services, which should be familiar, and which are using computer typesetting to get the information before the customer as quickly as possible. One is the Library Association's British Technology Index, another is the combined INSPEC service of the Institution of Electrical Engineers, and another is Chemical Abstracts. These three represent a whole range of similar systems.

An interesting pilot was the House of Commons index of last year, which I saw as the most significant contribution towards informing elected representatives of why they are there that the country has ever had. If penny-pinching prevents further examination of this service, then the British concept of democracy will have throttled itself in accepting the restriction of public spending as a worthwhile pursuit. It is essential that those who plan our future can quickly dis-
cover the knowledge that as voters we collectively have. Simply gaining more votes than an opponent does not transform a parliamentary candidate into an informed oracle. Democracy relies not only on the existence of information, but on the objective transmission of it. It has to be known to be instrumental.

If during the remainder of this paper I make statements about indexing that some of you may find quarrelsome, I hope you consider the particular contribution that indexing can make to an informed democracy—whether the process has been a human or a machine task. Countries that produce indexing are more democratic than those which do not. Even a Central Capability Unit needs organised guidelines to data before it divulges the benefits of its capability. It is equally important that we can assess the information that produces the judgments.

I see your members either as professional librarians whose daily function involves indexing a collection of some sort; or as freelancers compiling book indexes for authors or publishers. The librarian should have come to terms with the computer by now, and I hope he is well aware of the work that he needs to do in order to use the computer to ease his work. Strange, but true.

The other animal need fear nothing, because no one in their right minds will use a computer to index books, even where the book is computer typeset. All that would be produced would be a slab of 'over-indexing' of limited use, and taking up valuable collation. Extra book pages imply production difficulties with consequent increased cost. Such indexes would need meticulous editing and so the computer's contribution, although valuable, would be limited to providing an unpublished concordance—which the indexer could use as his starting point. It will be cheaper to have a human being read and compile from the start.

All this may seem an irrelevant and idealistic aside, but it introduces a thought which I will expand later—the relative unimportance of books, compared with the volume of fresh knowledge transmitted through journals and reports. Modern society is more threatened by technological change than by the mystiques of economics, or by the applications of various shades of political dogma, although I appreciate that there may be a connection. If the layman is to argue convincingly for rational change, he must be able to comprehend what is happening, and all the possibilities and alternatives.

When governments and corporations blind him with science, he must not stick his head in the sand, or wear dark glasses. He must be able to give as good as he gets. To do this he should not have to rely on esoteric data banks, to which only a meritocracy will have access. He should be provided with public guidance through available scientific and technical abstracting services. No power of authority should exist from a consensus of ignorance, particularly when power comes from the counting of hands. It is therefore imperative that the walls of this establishment transmit the message to all public librarians who pass between them. Unless public authorities realise the necessary curiosity of the common man, they are failing in the responsibility that has been given them. Lip service is paid to the position of fount of all knowledge conferred on libraries, but as far as science and technology is concerned, there is a terribly conspiracy bordering on censorship by default.

It is extraordinary that most public authorities allow complete freedom for the professional librarians to select book stocks, but periodicals are subjected to tight scrutiny by gathered bands of shopkeepers. If you do not believe me, investigate any large reading-room where you will quickly see that to library committees, the definition of a technical journal is one that I would call 'trade'.

Since the higher level technical journals are not normally read as newspapers, but are produced to record ideas and endeavours, it follows that they must be supported by abstracts and indexes. It also follows that
as disciplines cross-fertilize and as the flood grows, these tools must be as comprehensive and topical as possible. To me it is no non sequitur that such a target is impossible without computer-assistance at some stage, and fears of technical innovation allayed by the use of technology.

Since the computer is an indexing machine presenting the user with required information it has first to find it within itself. The processes of information retrieval that enable this to be done need to be specifically recognised and intellectually designed by a human brain. It is thus no accident that many terms of specific meaning within the computer world have been taken from the world of paper-based record. Any computer journal will be full of words like:

INDEXING LOOK-UP CROSS-REFERENCES
LIBRARY ROUTINES DIGIT PAGING
CUMULATIVE CONCORDANCE FILE
THESAURUS COMPLIER TRANSLATOR

Even if they appear far removed from the meanings we first learnt. Before you wince pedantically, consider the origins of the words BOOK, PAPER, PARCHMENT, and PALIMPSEST, which I used before.

In a computer-bound world, the indexer may have something to offer of value. For instance, before programs become systems, much documentation is required, and the techniques used to analyse such procedures are akin to indexing. Without going into detail, it should be easy to visualise the type of intellectual effort needed to ensure that the output provides the required information in the right place at the right time.

Long before the computer has handled it, a human mind has defined sources from which the program will compute, or retrieve, or take as direct data input. Meticulously designed routines apply to all forms of data processing, from invoicing to automatic translation; from payroll to process control. Without this work the computer is as empty as a ream of paper before it becomes a book, and in most cases the programmer cannot start work until these paths have been mapped for him.

Too many computer installations are ineffective because the systems designers were ignorant of indexing principles. I think the historical basis for this is that the accountant mentality dominated for too long, and that left- and right-hand pages of ledgers became translated into input and output. In a simplified payroll system, a worker has a number which is matched with a name held in store for printout; and with a department; and with a known basic wage. Hours of work can be keyed in, or optically read from special forms; and from this a program will compute gross pay, take all deductions, and present net pay. All this information has come from different sources—from direct input, from files, and from computation. In what I call a 'ledger-minded' system the data punched-in will include name, number, department, grade, basic pay, &c., &c. The result of this is high volume of data preparation and costs to match; vast error rate or complex verification processes; much time consumption. And all to save costs on good programs and time on the central processor. If it is essential to save in this area, one might ask why it is felt worth using a computer at all.

To return to indexing: in order to use a computer for this purpose, input preparation requires an extravagance of labour. Therefore the prime reason for compute application must be to handle vast quantities of data with a flexibility previously impossible, and the need for this must be dominated by a sound economic or strategic reason. Otherwise it is just not worth it—as the editors of European Chemical News have decided with the Lynch program. It didn't increase the circulation, nor bring any new advertising, and it increased the administrative overheads. To meet these problems would have meant increasing the cover price, or selling the index independently at the full economic rate, or sharing costs by involving other journals.
The organisation of information will be an imperfect process long after computers have evolved into something else. But as we progress towards devices operating more like human brains, the whole philosophy of indexing will evolve, since I cannot believe that my own brain stores and retrieves information by crude word processing.

As far as current work is concerned, the computer is applied to five main areas in documentation, some of which are more developed than others. These areas are compiling, editing, producing, transmitting, and interrogating. As human indexers, you only really cover the first of these, and the most developed computer work is probably in the last three areas, with editing a fairly recent application. The computer cannot compile anything until a human interface has presented it in a machine-recognisable form—which also implies selection. In other words, the computer is an invaluable ally in handling vast amounts of data to pre-selected routines, but it has not, at present, much contribution to make to selecting the data it is to handle. Any future efforts to question the social impertinence of the computer will emphasise this problem, but I don't foresee a potential tyranny, because even selection routines will need human anticipation and control.

I have indicated that some data processing installations have not been as efficient as justifiably expected, through failure to organise detail in a systematic way, so that all contingencies were foreseen after close investigation and adequate record. Thus one future of indexing lies here, both within the field of systems design, and, in future, in the application areas of on-line time-sharing systems. If the jargon worries you, the concept is readily grasped.

An on-line system operates in the same way as a telephone network (in fact, in many examples the Post Office telephone complex is used), except that at the other end are computer services. The user's terminal will assess any service for which the subscriber has paid—and these may be offered by many different tele-processing agencies. Each customer may think that he is the sole recipient of the service at any specific moment, but the system is capable of serving all customers at once, apparently instantaneously. This is what is meant by time-sharing.

One number may be dialled for mathematical computation, another for stock market information, or another for accounting or stock control routines. The future will see thousands of various services accessed as readily as today's plumber through the Yellow Pages, offering from transport timetables to guides to restaurants; from entertainment information to remote library references—and all from one's home, or from public kiosks.

It should not need imagination to see the indexer contributing both to the design and to the use of such systems—as well as using them himself to compile, store and edit his products, which are then input into a processing network.

There is no difference between this electronic image and the process of using books and indexes to produce more books and indexes—some of which are classified directories. You are probably aware that the telephone directories now use computers to set them in type, operating on output produced by computers which have processed entries into classes and alphabetical subsorts. To make it sound less sinister, it is probably all done by the same installation through different peripherals. Even so, we all keep our own little lists of telephone numbers—and probably always will.

I would like to finish with the computer by quoting from a paper presented at a Printing Industries of America conference in May, 1970; by an employee of *Time* magazine:—

"There was Mark I and ENIAC, UNIVAC and KWIC, UNITERM and PEEKABOO, POLYTERM and KIB,"
KAD, KOV, KWOC and KWAC,
CRAM, CHRIS and COMAC,
WALNUT, CYPRESS, SAGE, FLIP,
Magnavue and Microchip,
Filesearch, Filmsort, Filmorex,
TWX, PERT, Computer decks,
Multics, JOSS, ARPA, DARE,
BMEMS, PILOT, ATLAS, SHARE,
MEDIBANK and Solómon, WHIRL-WIND and SPAAC (sounds like "smack"),
An ERA, SNOBOL, WATS, even EDUAC.

The problem is that we are trying to use these new engines and their programs for very complex purposes—many of which we barely understand. Man is trying to embed machines into systems that are part-manual, part-automatic, part-electronic. There is no question that many investigations into new business and editorial opportunities have their roots in technology. Certainly the New York Times Information Bank would never be built without computerized microfilm and computer-aided indexing. Publishers would not enjoy subscribers in the millions without the electronic production of subscription fulfilment. Time Inc. got two new businesses that total over 100 million dollars as a by-product of its first subscription service system installed ten years ago, namely direct mail sale of books and sale of complex regional advertising.

I have twice hinted at certain book concepts which must be challengeable. During earlier discussion with the editor of your journal on a possible approach to this paper, it was suggested that I might comment on the quality of certain book indexes, with examples. However, I feel that some books should not have indexes, and state this heresy very firmly. I submit that a book can be structured so that it works without an index, and there are too many books that are badly structured; thus needing an index to sort out the mess. An example of a well-planned reference work succeeding without an index is the Oxford companion to music, although all the Oxford Companions rely on the cross-referring dictionary principle. Another is Brewer's Dictionary of phrase and fable, but again all dictionaries fit this. I believe the Roget's Thesaurus, although indexed, can be used without one. Any trade and commercial directory worthy of the name is self-indexing, the secret of retrieval lying in the user understanding the book. One librarian I worked with always used Whitaker's Almanack without needing the index; and it was drummed into all librarians never, never, never to use the index to the Dewey Decimal Classification in order to classify. And, in truth, a good classifier never needed it and a good librarian works from his shelves, his catalogue, or his experience.

It might worry you to know that the first man to compile an index, or concordance, to the Bible was burnt at the stake—because by doing so he had challenged the need for ecclesiastical guidance. A restrictive practice had been dented. I cannot help feeling that the church's anger had some justification. If you think about it, an index implies that the book can be thrust together in any order—scarcely an implication that should be made of a holy book in darker times. It is this implication that disturbs me. Just as some books could be designed to be more self-informative, there are others which though needing an index, become over-indexed.

It is sad that a very fine Victorian practice has been supplanted by the modern cruder index. I refer to the full contents pages carrying chapter synopses, which led to page references carrying marginal guides alongside the pertinent passages.

I remember a Times crossword clue that was an incomplete quotation from Byron's Don Juan, and the compiler had ensured that his 'light' remained hidden within the bushel of the Oxford dictionary of quotations. I was forced to read the poem in some depth. Of course, I could have waited for the solution the following day; but should I complain of the lack of indexing that allowed this experience? Are there not some books that
should be read, just as there are others that are to be used when needed?

I have an enjoyable social essay on railways by a respected author in which he pays tribute to his wife's index. It is a prime example of 'cottage industry' indexing. To the text covering 150 Penguin pages is appended fifteen pages of bad indexing. The book is predominantly British railway history, yet the index devotes one and a half of these fifteen pages to a subject heading RAILWAYS IN BRITAIN! To seek references to King's Cross, you have to be astute enough to refer to LONDON STATIONS. Since the author does not refer to Broad St., Liverpool St., Fenchurch St., Charing Cross, Victoria, Waterloo, Cannon St., or Blackfriars, I think the indexer is being too presumptuous in choosing a generic heading when a specific one is adequate. Faults are very easy to spot, but what is sad is that the single contents page is adequate enough to retrieve anything within 150 pages—particularly if one has already read the book. To add to a well-shaped effort a ten per cent index capacity seems to me a sacrilegious waste of time.

Meanwhile, John Prebble's absorbing account of the Tay Bridge disaster, The high girders, is not indexed—a state of affairs that aggravated me when seeking a point particular to this very paper. These comments lead me into the question which is not asked in polite society. Just why are books indexed? What policies exist within publishing houses to decree that some titles are treated and others not? Why do publishers spend much effort in producing house standardised packages from cover to cover—except that all their indexes are thrown together without any principle whatever?

I believe that every book should be closely analysed to decide whether an index is desired, or whether the author could structure his presentation more efficiently. If an index is necessary, then the type and level should be decided, and its production should be to house Specifications, at the very least. In other words, if the index is not presented as an extension of the author's work by the author, then it should be an integral part of the book production process. I realise that I am sticking my neck out, but I am not convinced that the tradition of farming indexing out on a smallholding basis has given indexing the consideration it deserves.

You, as indexers, may worry about the machine. As a librarian, whose living comes from utilising published information, I ask that if your level of indexing is capable of economic replacement by machine, then it ought to be so replaced. As librarians, we never work for ourselves, since we are always servants to some style of organisation.

Too much book-indexing is soulless. Many indexers do not appear to have read the book in their charge, and merely match words with page numbers. This is not indexing. It does not aid the reader, and it confuses the searcher. In some cases it might even destroy the author's concept. Although I believe that indexing should be based on identifiable standards, I include within this philosophy the requirement that the index of a book should be part of the complete texture. If you study the Oxford dictionary of quotations, the index shows itself as the product of overall editorship—not as the work of isolated cottage labour.

Before a final comment on the rôle of the book in our future society, I would like to throw in an aside on indexing as a necessary craft—and one justifiable reason for indexes for the sake of indexing. In organising our own thoughts, studies, curiosity, or collections—to index is to learn. I think that if I were to index my own collection of science fiction short stories, I should not publish it. It is the duty of others with similar passions to undergo the same experience for themselves.

I will close the paper with my contribution to the fashionable debate on the future of the book. Everything I have read by experts I find wide of the mark that I gauge. One side forecasts books to be replaced by electronic or optical media; while the other side can see nothing other than the book as the
only convenient form of communication, most suited to the time and privacy of the individual. Two technologies are trading blows. The question should be, what type of presentation is suitable for book format? If fiction is a decayed fashion, how soon will it be replaced by a new imaginative art form based on, say, video cassettes? If directories are costly to compile, tortuous to typeset, expensive to distribute, expensive to shelve, diabolical to use, and difficult to rely on—at what stage will micro or digital formats service a happier market? Is poetry better borne by oral tradition than by the artificial formalities of the printed page? Will telephone directories be part of the hardware, rather than ergonomically bad, closely printed, floppy pages thrust periodically in bumper bundles through each door? Should not translating or grammatical dictionaries be provided through interactive online services, since it is constantly argued that language is a living organism changing daily? Printed dictionaries freeze a language, while today’s electric global village needs greater flexibility in communication. If we are to class both novels and telephone directories as ‘books’, we are using a technical tradition to simplify a complex landscape.

Words are barriers where they prevent the transfer of a concept. Otherwise painting, sculpture, photography, and music would never have needed to exist. There will be a ‘Newspeak’, not as Orwell envisaged it, but so that communication can be more symbolic, more graphic, more cyphered, more conceptual. At the moment we are lacking the language, and a latter-day Chaucer will emerge—and he might even be a disciple of Marshall McLuhan.

Chaucer was a contemporary of Gutenberg—and McLuhan is ours as we move towards this perplexing future of the moving finger. I will close with my favourite quotation on futurology and the recalcitrance of man, aptly closing the chapter called ‘Theories’, in Crotchet Castle:

‘The schemes for the world’s regeneration evaporated in a tumult of voices.’

References


SUMMARY OF THE DISCUSSION ON MR. GEE’S PAPER

In the discussion which followed, the audience appeared reluctant to take up the big issues which the speaker had raised, nor did they attempt to rival him in prophecy about future developments.

The discussion opened with Mr. Neil Fisk objecting to Mr. Gee’s criticisms to the index of Roget’s Thesaurus. Mr. Fisk thought the index admirable and explained that he often used it more effectively by looking up two words approximating to the one he was trying to find an alternative for, and when he saw the same number under both, this generally referred him to the word he wanted. The speaker explained that he was a great browser in books and he found it
difficult to consult Roget's great work at speed, but he always wanted to stay longer, so he found the index a great time consumer. Also when he first used Roget, he only possessed volume one of the Everyman edition and the index was in volume two; so he learnt how to find his way around without using the index. But he considered it such a well-structured book because every entry referred you to some other entry.

Mr. Molesworth Roberts asked Mr. Gee if he considered that many page entries under one heading indicated bad indexing. Mr. Gee said that many books could be much better structured than they were, thus less reliance could then be placed on the index to draw subjects together. Both questioner and speaker agreed that biographies demanded close indexing and Mr. Gee again regretted the passing of the Victorian habit of extended chapter headings and margin entries, which he thought a great help to the reader.

Mr. Raper pointed out that the area of subject matter in a technical index was voluminous, and instant rapid retrieval was better than browsing. Mr. Gee felt that personal browsing to find information for oneself was a different matter to finding facts for others and reiterated his feeling that too many books have indexes they don’t need and many others which need them don’t have them.

Mr. Gee also expressed concern at the pitifully small collection of journals on science and technology held by most public libraries, and thought they should spend more on these subjects. He also was very well aware of the mainly inadequate indexes (if any at all) produced for these journals, their usefulness being thereby greatly lessened.

Mr. Raper suggested that as knowledge was doubling every eight to ten years it was essential that libraries subscribed to more abstracting services. A response from the audience complained of the high cost of such services, exemplified by Chemical Abstracts at £400 per year. Mr. Gee felt that such expenses were essential to an efficient information service, but agreed that he could not justify CA although he spent £1,000 per annum on other abstracting services.

To the criticism from the audience that such scientific and technical journals would not be used sufficiently in public libraries to justify their cost, Mr. Gee pointed out that the same arguments were levelled against the provision of books when public libraries first came into being. He felt it was the fundamental task of libraries to educate and inform. This was particularly important at the present time as there was such a widespread ignorance of science and technology which was not improved, he felt, by the almost complete absence of information on these subjects in the national newspapers, the Financial Times excluded.

To a question put by Mr. Norman Knight on the necessity for the I.E.E. Electronics Abstracts to have a multiplicity of separate indexes, Mr. Gee pointed out the need to avoid confusion in a single-sequence dictionary index when covering 12 monthly collections of highly specific subjects; authors; institutions; companies; trade marks; and patents—particularly when each separate index was compiled by separate indexers. This problem implied computer assistance. Many visually attractive indexes were now being produced by computers with the use of both upper and lower case letters (previously held to be impossible) and also computer typesetting.

Mr. Gee ended by complaining that he thought that, for the most part, the state of British abstracting was abysmal, though there were one or two bright patches, especially the abstracting services of the Institute of Physics. It was also pointed out that the American Chemical Society had signed an agreement with the German and British chemical societies to provide information.

Nevertheless, Mr. Gee concluded that publishing in Britain is not an information-conscious industry; it is a selling industry, and the source of his salary.