

Stone Age thinking at the speed of light – AACR2, MARC, and other dinosaurs¹

JOHN JOLLIFFE²

Transcribed from a cassette tape recording and annotated by David Helliwell with the help of Geoffrey Neate. Thanks to Terry Belanger, Kay Guiles, and Robin Alston for valuable information on the persons and events mentioned in the lecture. March 2007.

I welcomed the opportunity of preparing this talk before I welcome, time permitting, the opportunity of attempting to crystallise my ideas on some topic which has been growing in importance in my mind over a period. I'm not the sort of performer who dons cap and bells and is happy to give for the fiftieth time a recitation of the Seven Ages of Librarianship. After suggesting this talk to Terry³, and then seeing the other speakers here included in his programme, I realise that I might have perhaps better chosen a different topic, another one on which I must soon crystallise some thoughts, and that is on certain aspects of book pricing and distribution on the Continent during the sixteenth century, but I'm afraid you'll have to wait for that one.

The conversation I had with Terry last fall, which led to the idea of giving the talk, had been on matters nearer the present idea, and so I naturally thought of this as being the one that he would like me to dilate upon. I must apologise in advance for some comments which may be not generally understood because they involve technicalities of computer processing or of programming. Such comments will, however, tend to reinforce a proposition which has been commonly accepted since the first days of using computers in libraries, namely that librarians should know more about computers, not simply from the outside, knowing the uses to which they are put, but from the *inside*, knowing by using them at practical level what they are capable of being used *for*.

¹ Book Arts Press lecture 116, 17 March 1983. This series of occasional public lectures was inaugurated at Columbia University School of Library Service in 1972. It continues under the auspices of the Rare Book School at the University of Virginia as "Rare Book School lectures". Lecture 500 will be delivered on 16 July 2007 by James Green of the Library Company of Philadelphia.

² John William Jolliffe (1929-1985) was Assistant Keeper in the British Museum's Department of Printed Books, 1955-1970, Keeper of Catalogues in the Bodleian Library, 1970-1982, and Bodley's Librarian, 1982-1985. Obituary by R.J. Roberts in *The Bodleian Library record* 12:1(1985), 1-2. See also note 11 below.

³ Terry Belanger was then Assistant Dean, School of Library Service, Columbia University, and Director of the Book Arts Press, the lecture's sponsor. He is currently University Professor, Honorary Curator of Special Collections, and Director, Rare Book School, at the University of Virginia.

A very long time ago, I was introduced to a statement by Ada, the Lady Lovelace, daughter of the poet Byron, mathematician, and supporter of Charles Babbage, the inventor of modern computer programming.⁴ Referring to the machine which Babbage was designing and building with government support over about twenty years, and without success, she said, “It will do whatever we know how to do.”⁵ Now the meat in this proposition lies in the words “whatever we know how”, and I’ve found it to be true in nearly twenty years of work with computers. In this time I have also met remarkably few librarians who knew either the statement, or its truth. I have however met many who have worked with computerised library systems who are unaware of this truth and of the obligation it lays on us, not on the machines, for the limitations of the systems which we design and implement.

Looking back to the early 1960s, it is possible to discern a simultaneous development of two features, which especially when they have interacted, have begun to change the practice of libraries. On the one hand there has been the continuous development of the most sophisticated and flexible instrument ever designed by the mind of man – the computer – and on the other there has been the introduction into library practice of ever more restricting codes, and ever less flexible standards. This is ironic.

A further irony is that in recent years it has sometimes been said in justification of a new restriction that this is because of the requirements of the computer. I will put forward a thesis, which you may be relieved to hear I will not attempt to justify word by word, that the thinking behind the use of computers in libraries is out of date, and possibly has always been so, and that the opportunities for newer modes of thought and newer methods offered by technological advances have largely been ignored. It would be possible to argue, as some computer specialists are now arguing, that by 1965, the computer structures based on the Von Neumann logic of the 1940s, and on the limitations imposed by the earliest component of the computers, were out of date. I shall return to this briefly later on.

⁴ Augusta Ada, 1815-1852, daughter of the poet Byron. In 1835 she married William King who later became the first Earl of Lovelace. She financed Babbage’s work, and is credited with having written the world’s first computer program.

⁵ Her actual words were: “The Analytical Engine has no pretensions whatever to *originate* anything. It can do whatever we *know how to order* it to perform. It can *follow* analysis; but it has no power of *anticipating* any analytical relations or truths. Its province is to assist us in making *available* what we are already acquainted with.” Note G in L. F. Menabrea: *Sketch of the analytical engine invented by Charles Babbage ... with notes upon the memoir by the translator Ada Augusta, Countess of Lovelace* (in Bibliothèque Universelle de Genève, 82, October 1842).

For the moment I want to remain with library system designers, rather than computer designers. My first contact with computers was in 1964. At that time I was working in the library of the British Museum⁶, and the head of the library, Robert Wilson, had been asked by Sir Frank Francis, who was then the Director of the Museum, “What are you doing about this new idea of using computers in libraries?” So I was given fourth refusal of the question, “Would you like to find out about this idea of using computers in libraries?” I said “yes”, and the story should go “I have never looked back.”

In the following year, together with a colleague from the Library, I came to the United States for the first time and for a period of three weeks the two of us were almost in daily contact with those in the Library of Congress who were planning the MARC project. I do have a longish memory, although I hope I’m not yet a fossil. In those days Snyder was not Henry [laughter] and ESTC North America, but Sam and the LC Systems Office.⁷ He was Mrs Avram’s⁸ colleague and superior. What has been fossilised from that time is MARC itself. Then they were planning an experimental distribution of cataloguing data on magnetic tape to, I seem to remember, only fourteen libraries, and for a limited period of time, perhaps six months.

The inevitable happened. Those first thoughts, subject in principle to revision and improvement in the light of experience, are still with us. The libraries participating in the experiment, which had invested so much effort in programming, and which had altered their systems to depend on the continued availability of cataloguing in the experimental form, were naturally reluctant to encourage any radical modification of the MARC structure. In consequence, the experimental nature of the MARC 1 project

⁶ This library, with several other national collections, became the national library under the terms of The British Library Act of 1972. The BL came into operation from 1 July 1973.

⁷ In the 1980s, the audience would have taken the name “Snyder” to refer to Henry Snyder, who had been appointed director of the ESTC in North America in 1978. But in the 1960s, the name referred to Samuel Snyder, who was appointed Information Systems Specialist at the Library of Congress in 1964. His task was to draw together a number of different automation initiatives in North American universities with the aim of producing a standard approach to the formation of a national pool of authoritative bibliographical data. The LC’s “Office of the Information Systems Specialist” had been renamed “Information Systems Office” in 1965. See K.M. Spicher: *the development of the MARC format*, 78 (in *Cataloging and classification standards and rules*, The Haworth Press Inc., 1996, 75-90).

⁸ Henriette Davidson Avram is the creator of MARC. She was born in New York City on 7 October 1919, and after working for seven years at the National Security Agency as a computer programmer and data analyst, she joined the Library of Congress in 1965, where she was given the task of designing and implementing electronic cataloguing. The project was completed in 1968, and MARC became a US national standard in 1971 and an international standard in 1973. During her 26 years of service in the Library of Congress, Avram was responsible for most automation and networking functions. She retired in 1992, and died on 22 April 2006 at the age of 86. She is the author of *MARC, its history and implications* (Library of Congress, 1975).

really did not come to any conclusion. Certain features, which even in their design phase in 1965 were acknowledged to be weaknesses, are still with us. Effort was concentrated on modern books, indeed on modern books in English, and all the arguments of myself and my colleague for a wider specification, within which our problems at the Museum of books of all dates and in all languages might be accommodated were listened to, but set aside in favour of the experimental format, and the coding on to which it was later necessary to graft, and not always easily, further features to accommodate non-English and non-modern books. Not all the necessary features have yet been added.

At that time, generality was applied to the MARC record itself. It was to be independent of the medium of transmission. And so it received the character-string format, even for numerical references within the record, which is still a feature of the international standard, though to the best of my knowledge no-one ever wished to transmit MARC records over telegraph lines to teleprinting equipment.

Another relic of this early preoccupation with, in effect, validation is the occurrence of field-terminator characters, as well as machine-generated pointers to fields, which give the length *including* the terminator. Thus to get the data from the field, one has to start by subtracting one from the length. Belt-and-braces indeed, both supplied by machine!

And how did this happen? The logic of circumstance and the logic of need only partly accounted for it. One must remember that in those early days, the only solid input from the library side was from Mrs Markuson.⁹ Perhaps I can depart to a small anecdote about Mrs Markuson here, which is not intended to disparage her.

During that three weeks we were in Washington, there met an organisation called COLA – and I don't know if it still exists. It stood for Committee on Library Automation, and it had apparently started as being a talking group in the corridors and bars of ALA, being those people who were already starting to try to use computers in libraries.¹⁰ By the time we met it had become a little more formalised. It had a chairman, though it

⁹ Barbara Evans Markuson was an assistant of Samuel Snyder in the "Office of the Information Systems Specialist".

¹⁰ COLA was the re-named "Clinic on Library Applications of Data Processing" whose second meeting was held in April 1964 at the University of Illinois in Urbana. "COLA [...] became a discussion group within ISAD [Information Science and Automation Division of the American Library Association] in 1970. Five years later, the group changed its name to the Library Automation Discussion Group, and in 1981 it merged with the MARC Users Discussion Group to become the Library and Information Technology Discussion Group. This group lasted until 1984, when it was disbanded because of low attendance." <http://www.lita.org/ala/lita/aboutlita/org/1st25years.cfm>.

did not yet have an agenda and minutes, and we were asked if we would like to sit in as observers, which we did. And then one evening Mrs Markuson provided entertainment at home for all the group, including ourselves. When we arrived, she met us by saying, “Well, after the meal you and Sandy Cain¹¹ can tell us about your plans for the British Museum.”

We were saved from this in effect by the fact that what was served as wine with the meal was sherry, and Sandy Cain and I, finding it was sherry, took it very slowly. Others drank it as if it was Algerian plonk [guffaws from the audience] and after dinner when we came to give our talk, I think we were probably the only two in the room who were still awake.

Now, as I say, in those days there was only Mrs Markuson who was from the library side, and the dominant people had come from computing, and as far as I remember, computing of a very strange form, in code-breaking, here and there. They tended to bring with them some of their previous experience, but they also saw the library world in concrete terms. And what they saw as the most obvious manifestation of library information was, if you’ll pardon the expression, the 5" x 3" card.

Cards were still in vogue in computer systems. You’ve only to look at films of the 1960s to see that whenever a computer or a retrieval system has to be demonstrated, what you actually see is a punch-card collating machine from which someone draws the one card which will indicate whatever it is.

In those days the terminal as an input device did not exist. Key to magnetic medium input devices for off-line data capture were in their infancy, and since they’d been manufactured by the *typewriter* division of IBM they were naturally incompatible with the *computer* division of IBM (which caused a certain amount of difficulty here and there). The punch-card was almost everywhere dominant. And so it was natural first to think

¹¹ Alexander Matthieson Cain was Assistant Keeper in the British Museum’s Department of Printed Books, 1954-1966. He and Jolliffe, “two of the most able of the younger Assistant Keepers”, were set to work on the task of library automation in 1964, and visited the U.S.A. in autumn 1965 to investigate the progress that had been made there. Cain left for a post in the U.S.A. in autumn 1966, and Jolliffe completed their confidential *Report on the feasibility of using automatic data processing in the British Museum, principally in the Department of Printed Books* alone in April, 1967. See P.R. Harris: *A history of the British Museum Library, 1753-1973* (London: The British Library, 1998), 636. According to Robin Alston, “ In the history of library automation [this report] is crucially important: every bit as important as the discussions which took place at the Library of Congress in 1964 regarding the creation of a computer tagging system for producing computer-based catalogues, the practical result of which was the development of MARC, now the universally accepted tagging code for library catalogues.” See Robin Alston: *The Eighteenth-century Short Title Catalogue – a personal history to 1989* (www.r-alston.co.uk/estc.htm).

in fixed field ways, and the leader of the MARC record is still a punch-card.

It was natural also in those days to think of asking library staff to prepare their records for the machine in block capitals, on 80-column wide coding sheets. IBM had not yet admitted the possibility of lower-case letters, and never has adopted the international character-set standard, which is built round the common alphabet.¹² But the library card gave the model for the order of presentation and tagging of data elements, and accounts for the curious continuing presence of both fields and sub-fields in the record, which actually requires two modes of access within the same record to certain data elements. You use a pointer to find the field, and then you have to do a character-by-character search through the field to find a sub-field. Why not do one or the other? Why do both?

The MARC record contains these subfields because on certain lines on the card you will have, for example, an imprint, and there it all is – place, publisher, date, and size, so the line is a field, but nevertheless we want to get at these individual things, so we chop it up within the line. The notion of having several fields on the same line was impossible of conception.

The MARC record is not totally unchanged since the beginning. The separation of the data from the pointers, so that the pointers refer to the area of the data only, and not to an area relative to the start of the record, is a modification which was intended to permit changes of the record by addition or deletion of data without the necessity of changing all the addresses in all the pointers. But the pointers still appear in numerically ascending order of tag, but there is no necessity for this. And in spite of the generality of the pointer system, the data fields are also concatenated in tag order. Even where it is possible, no two pointers point to the same data. On the typed card, if “London” is needed twice it is typed twice because it is in two different places. In the MARC record, it would be perfectly simple to have two pointers pointing to the one word “London”.

¹² It has to be remembered that Jolliffe is talking in 1983 of the situation in the mid-1960s. At the latter time IBM was still using BCDIC (Binary Coded Decimal Interchange Code), a 6-bit character set with 64 values and which allowed only upper case versions of the alphabetic characters. This was enlarged in 1970 to EBCDIC (Extended Binary Coded Decimal Interchange Code), an 8-bit set with 256 values and including lower as well as upper case letters. Both sets had inconvenient gaps in numerical values between the three sequences A to I, J to R, and S to Z. EBCDIC was the standard used by the widespread IBM 360 range and continued to be used by the company even after ASCII (American Standard Code for Information Interchange) had become ISO 646 in 1972. When this talk was given IBM had just adopted ASCII for the IBM PC which then ushered in the modern world of personal computers.

For some data elements there is still no provision, for example, bibliographical format, bibliographical collation, a reference to the bibliography in lieu of a standard book number, for example a reference to STC.¹³ But here I start to tread on different ground. It might be unreasonable to expect MARC to provide for elements absent from AACR, just as it may be unreasonable to expect MARC *not* to include elements in AACR for which there is no logical justification when applied to certain kinds of book – height of copy, for example.

But before coming to AACR, a brief diversion into the ISBD and the ISBN.

I remember the time when the ISBD was no more than a gleam in Michael Gorman's¹⁴ eye. I remember why it was proposed. The notion was that with the growth of MARC-based national bibliographic centres, it would be advantageous if bibliographic records from countries which had a national bibliography but did not yet produce MARC records on magnetic tape, if their records could be put in such a form that they could be typed into the computer in the more advanced countries, and put through a single program called an automatic format recognition program, which would turn the records into fully tagged and structured MARC records.

The idea was not absurd. The implementation has been.

Basically, only those advanced countries which produce their national cataloguing in MARC form have adopted the ISBD. This is still not quite absurd. What is absurd is that those countries, which because they had their records in machine-readable form, could print them out in any way they wanted – backwards, putting “God save the Queen”, twelve asterisks, or the text of the Gettysburg address if they so wished as the punctuation between title and imprint – have adopted the barbarous and illiterate punctuation of the ISBD.

¹³ Pollard & Redgrave's *Short-Title Catalogue*. By this time, the STC revision (by W. A. Jackson and F. S. Ferguson) was in progress, and the second volume (I-Z) had been published in 1976; the first volume (A-H) was published in 1986, and the index volume in 1991. The full title of the work is *A short-title catalogue of books printed in England, Scotland, & Ireland and of English books printed abroad, 1475-1640* (London: Bibliographical Society).

¹⁴ From 1966 to 1977 Michael Gorman was successively Head of Cataloguing at the British National Bibliography, a member of the British Library Planning Secretariat, and Head of the Office of Bibliographic Standards in the British Library. He is the first editor of *Anglo-American cataloguing rules, second edition* (1978) and of the revision of that work (1988). He is the author of *The concise AACR2*, 3rd edition (1999). He has been Dean of Library Services at the Henry Madden Library, California State University Fresno since 1988.

Because they did not realise the real independence of the input record, the processing record, and the output record, the intended input record has become the normal output record! Not a great amount of straight thinking there.

The ISBD also illustrates another feature of some recent progress towards standardisation: the tendency for one standard to prejudge an issue not yet itself standardised. In this case the ISBD depends essentially on the acceptance of AACR as an international cataloguing standard. Put another way, this means if you don't like AACR, then you won't want the ISBD.

Another example of this sort of prejudicing of issues, perhaps a little clouded by real international politics, comes from the potentially unexciting field of character sets for bibliographic use.

There is an international standard for the Roman alphabet, and the characters such as numerals and punctuation commonly used with it. This standard is virtually identical with the American ASCII code. No-one has ever asked, and no-one has ever thought another order would be possible; but no-one has ever asked why the letters in this code are in alphabetical order – they are, it seemed a reasonable way to do it. In both these codes, the numerical values assigned to letters are in two blocks, one for upper case letters, and the other, with higher numbers, for lower case letters. No-one ever thought why you should do this, in that particular order, but it was done. A random choice, if you like.

Within each block the normal alphabetical order is followed with each number one greater than its predecessor. Thus capital “A” has the value 65, capital “B” 66, and so on, while the lower case “a” has 97, lower case “b” 98, and so on.¹⁵ This makes sorting of English and most western European languages quite straightforward. But not all – Spanish, Welsh, Danish, and other small countries of that sort cause some difficulty.

A considerable amount of work has been done over the years to provide additional standard sets for other groups of characters, the Greek alphabet, for instance, a further roman set containing common accents, and characters such as the “æ” and “œ” digraphs. One proposal came forward at the International Standards Organisation for an extended Cyrillic character set.

¹⁵ See note 10.

Now the Cyrillic alphabet presents a problem. It has more characters than the roman alphabet, so that it is difficult to squeeze it into a set of numbers of the same size. In addition there are in use in Bulgarian, Macedonian and Serbo-Croat some characters which do not appear in the Cyrillic alphabet used in Russia. There is also the need to provide for a certain number of characters in the Cyrillic alphabet which were discontinued after the revolution but which still appear in nineteenth-century Russian book titles.

The extended set was intended to include these additional characters for Bulgarian, Macedonian and pre-Revolutionary Cyrillic, and not the characters which the Russians use when writing in Cyrillic the non-Slavonic languages of Soviet Asia – a further set would be necessary for that in due course. Now this character set had been prepared by the Russians, and was based on their own domestic standard, called GOST (GOST is the name of their standards organisation)¹⁶. The International Standards Organisation accepted it on the nod, and it was circulated to national standards organisations for voting – ISO has a democratic structure. Majority voting counts. Indonesia and Thailand together can nullify the United Kingdom and the United States. For some reason best known in the Library of Congress, the United States representatives of ISO acquiesced in this procedure. Since it was in the euphoric phase just after the Helsinki agreement¹⁷, perhaps no-one wanted to be beastly to the Russians.

The voting was interesting. Of the sixty or so national organisations who voted, only one negative vote was recorded – that, I am proud to say, of the British. We voted against it on these grounds. First, that it claimed to be an extension of a set which had not itself been circulated as a standard. Second, that the block of numbers assigned to lower case letters was lower than that assigned to upper case letters, thereby reversing the procedure in all previous sets. And thirdly, that the alphabetical order normally assigned to the letters in the Cyrillic alphabet had not been followed. Now since we had the last two of these objections also to the parent GOST set, we did not wish to let this slide into acceptance undebated, because then the parent set would then have had to go through on the nod because the extended set had already been accepted.

¹⁶ GOST (ГОСТ) is an acronym for *gosudarstvennyy standart* (государственный стандарт), which means “state standard.”

¹⁷ The “Helsinki Accords” were the “Final Act” of the Conference on Security and Cooperation in Europe, held in Helsinki in December 1975. The principles it formulated for relations between sovereign states did much to reduce the tension of the Cold War.

Now the committee of which I am chairman has been subjected to considerable pressure by the standardisation bureaucracy to change our vote, for the sake of international agreement, harmony, and solidarity. Only one of the bureaucrats we have dealt with has given any weight to our intellectual arguments. We are nevertheless going ahead with the production of a British standard for Russian Cyrillic, which we shall propose to ISO as a model.

A last word on this to illustrate part of the problem.

The first three characters of the Cyrillic alphabet are, as it were, “A”, “B”, and “V”.¹⁸ The “B” doesn’t look like anything, but “A” and “V” look like “A” and “B”. The GOST standard assimilates where possible Cyrillic letters to the values given to similar shapes in the roman alphabet. That is, 65 goes to their “A”, 66 to “V” (because it looks like “B”), and “B” – their “B” – to a much higher value because at that stage something like Q, for which there is no equivalent. Thus sorting is no longer simple. Because the sorting order values – alphabetical order – differ from the numerical code values. And in objecting, we are trying to ensure that in England anyway, libraries don’t have to go off into another sorting procedure when they hit Cyrillic.

Another standard which has had much more acceptance, certainly in Britain, is the International Standard Book Number.

Now although the spelling mistake was not an invention of the early printers, it seems reasonable to say that they and their successors have institutionalised it and disseminated it in a way that the mediaeval scribes and their predecessors could not have hoped to emulate.

Now the designers of the standard book number did have error and its avoidance in their minds – why else did they produce the elaborate means of supplying the check digit at the end? Yet they seem to have had total confidence in printers.

Any bibliographer could have warned of the practices which were likely to arise in the use of the standard book number. Indeed, any bibliographer might have addressed the problem, what is a standard book number. This is not just an exercise in philosophy. I would expect most librarians to say, that the standard book number is a number printed in the book which is a unique identifier for the edition of which that book forms part. Now please don’t question if I’m right.

¹⁸ A, Б, В.

Now the standard book numbering agency in London, and the producers of British National bibliography as well, hold that the standard book number is a number assigned to an edition whether printed in it or not, and regardless of any apparent standard book number (differing from the assigned number) which is actually printed in the book.

From this conflict of definitions, and from certain publishing practices, for instance, perhaps the sensible one of providing a different standard book number for a hardback and a paperback edition, the less sensible one of sometimes providing the same standard book number for all four volumes of a four-volume work, sometimes insisting on providing four different numbers, sometimes providing a new standard book number for a revised edition and sometimes sticking to the old number, sometimes even using a standard book number for a new work because the work to which it was originally assigned is out of print and therefore the number becomes available – all these things exist, and this through publishing practice. But of course the printer has the last laugh.

We tried to urge the British National Bibliography to have an indicator – and there is provision for indicators in the MARC record – an indicator against the standard book number which will show whether it was actually present in the book. Or perhaps a range of indicators, one to show that the standard book number was printed on the back of the title-page, where one expects to find it, one to say it's on the dust jacket only, and a third one to say this is a number that has been assigned by the agency but doesn't appear in the book.

But it's not even quite as simple as that. I have on my shelves two OUP books in which I can see the standard book number, but the standard book number in normal form is nowhere printed in it. And what OUP have done there is take the numeration part of the standard book number, leaving off the check digit, and leaving off their own prefix, and use that number in the signatures. Now, is the standard book number in that book or not? As I say, it's the sort of question a bibliographer could reasonably address. We also urged them, as part of descriptive cataloguing, to record as standard book number what the printer had actually printed, regardless of any inherent errors. This latter they now do, though this caused them a certain amount of difficulty.

At one stage we wanted to test their card service, and we took two MARC tapes a month apart, and by applying a randomising procedure to it we printed out from these two MARC tapes two hundred standard book

numbers from each, and we printed them out on the teletype and we simply tore the paper off the teletype, and with a covering letter saying “Please supply these cards” sent it off to the British National Bibliography.

In each case we had a letter back accompanying 198 cards saying “Please check the other two”. Well, untouched by human hand! At no stage had we transcribed what was on the MARC tape. It was conceivable there was an error in the teleprinter, but it didn’t seem very likely, and being reasonably well placed in Bodley with the bulk of British publications coming into the Library, we shot it off and looked at the books themselves, and sure enough, what was on the teleprinter, what indeed was on the MARC record, was what was on the book, and it was true the check digit didn’t check. So although they had forced these things somehow, somehow, through their validation procedure to get the record sitting on the file, it was not possible similarly to force through the validation procedure any string which would actually match it, from which I concluded that perhaps one per cent of the MARC records were never going to be available to anyone.

Now, AACR is currently undergoing revision. What number will emerge is hard to say – a series whose first two terms are 67 and 2 can have almost any number as its third term. After the fuss about the cost of change from 67 to 2, one is surprised to find that undaunted, they are still working towards perfection. [guffaws from the audience throughout this section]

One of the more honourable scars I bear comes from a rebuke by John Rather¹⁹ for saying that the base of all cataloguing systems is economic. I meant that cataloguing systems were a compromise between what the library can invest in terms of labour for creation and maintenance of catalogues, and the utility of the catalogue to the user of the library. He objected to the notion that cataloguing rules were anything other than some branch of philosophy. Or was it theology? Certainly his party were in control, even though they shifted their ground over the years.

AACR2 abandoned one of the tenets of the Paris conference on cataloguing principles,²⁰ one of the few tenets which I thought was

¹⁹ At the time of Jolliffe’s talk (in 1983) John Carson Rather (born 1920) was Assistant Chief of the Descriptive Cataloging Division in the Library of Congress. In the late 1960s he became the Specialist in the Technical Processes Research Office. He became Chief around 1972, and was in that position when he retired from the Library a few years later.

²⁰ The International Conference on Cataloguing Principles, organized by IFLA, was held in the Unesco Conference Building in Paris, 9-18 October 1961, under the presidency of Sir Frank Francis, Director and Principal Librarian of the British Museum. It was attended by delegations from 53 countries, and observers from another 22. The outcome was a statement of 12 principles known as the “Paris

actually an advance on Panizzi, namely that all the works of an author, under whatever name they might be published, should be brought together under a single heading.²¹ Not any more. The title-page has become dominant, and the cataloguer, one assumes, is less and less expected to read the work, more and more to look at it and describe what he sees. Now valuable as objectivity may be, this seems to me a negation of what people can do better than machines.

Soon after that first trip to the United States in 1966, my colleague and I gave a series of talks to our colleagues at the Museum about computers and how they might be used in libraries, and especially how they might be used in the British Museum Library. It was a disastrous exercise in public relations. We were abused for selling out, for abandoning scholarship and scholarly preoccupations. We were also asked what we saw as the role of people in our machine-ruled Utopia.

We replied that the object of using computers was to get librarians back on the reference desk, where their ability to read, remember, and associate would be unchallenged by machine, and away from repetitious and mechanical transcription and recording which were properly the province of machines. This prospect did not go down well with the cataloguers either. [loud laughter from the audience].

I still think this is right, and I think it is sad that progress, ISBD, MARC, and AACR, have been going determinedly in the opposite direction.

There is a basic question about cataloguing which has been asked more and more over the past decade, and partly I am glad to say under economic pressure, and to which no reply has yet been given which justifies our present standard practice: what is the function of the catalogue? The question is fogged by the assumption on the part of the providers of cataloguing that they therefore know about catalogues.

It used to be said in contrasting the cataloguing practice of the British National Bibliography and the British Museum that the former, working

principles”, which have served as the basis for all national cataloguing rules ever since, including AACR2. The report was originally published by IFLA in 1963, and was photolithographically reproduced with an additional introduction in 1981.

²¹ The first major English-language cataloguing code was formulated by Sir Anthony Panizzi for the British Museum catalogue. His ninety-one rules were “sanctioned by the Trustees on the 13th of July 1839”, and underwent continual revision until 1936. They were published in the introduction to his *Catalogue of printed books in the British Museum, volume I* [no more published] (1841) in a section entitled “Rules for the Compilation of the Catalogue” ([v]-ix). Jolliffe’s understanding of Panizzi’s position does not seem to accord with rule XI, which states: “Works of authors who change their name or add to it a second, after having begun to publish under the first, to be entered under the first name, noticing any alteration which may have subsequently taken place.”

without books for comparison or files to maintain, catalogued each book as though it was the only book in the world, while the latter catalogued each book in relation to all the other books that had ever existed, whether it possessed them or not, Now both positions are a little extreme, and both, it could be said, fail to afford any primacy of interest to the non-librarian user. Unlike Panizzi, whose resolution of difficult cataloguing decisions was generally achieved by invoking a pragmatic utility to the user, which is why cataloguers find them so difficult to work with – there doesn't seem to be any theory behind it ... [laughter from the audience].

One of the examples there is the heading which is now abandoned called “Periodical Publications”, which Panizzi had. Panizzi said “Well, you treat institutions as authors, so that if you have a periodical published by an institution then you put it under the heading appropriate to the institution.”²²

But then you will come across periodicals which don't appear to have authors, *The New York Times*, or something like that. So what do you do with those? Well, you could scatter them throughout the catalogue by title. But he said it would be simpler to bring them together in one place, so that there was a large chunk called “Periodical Publications” where you would have not merely those periodical publications without authors, but also references to those which *did* have authors.²³ This was not really good enough. But it actually worked, and for certain types of bibliographic enquiry, which had certainly not been foreseen by Panizzi, for example the early bibliography of Norway or Denmark, by looking up the heading “Periodical Publications – Christiania”, you would find all the early publications of Norway. But I'm afraid reason had to prevail, and it is now not possible to discover in a single place the range of periodicals which the British Library has.

Anyway, the catalogue.

One of the functions of the catalogue in most libraries is to put the user in contact with a book. Several people have advanced the notion that if this is the *prime* function, then it can be performed with less cataloguing

²² Panizzi's rule XXXIII states that when the author's name does not appear “on the title”, if the work concerns or is addressed to a named person, that name should be taken as the heading. Rule XXXIV goes on to say: “When no such name of a person appears, then that of any assembly, corporate body, society, board, party, sect, or denomination appearing on the title to be preferred ...”.

²³ Rule LXXXI. The heading “Periodical Publications” was to embrace “reviews, magazines, newspapers, journals, gazettes, annuals, and all works of a similar nature, in whatever language and under whatever denomination they may be published.”

detail than we customarily and standardly provide, that is to say, with short records.

Peter Lewis, the director of the British Library's Bibliographic Services Division, which produces the British National Bibliography, is now veering in this direction, though not yet on paper. Farewell logical and philosophical purity! He now finds that it is impossible to contemplate a reduction from the seventeen weeks average which elapse between the publication of a British work and the provision of the full MARC record. So he is testing out the idea that he can improve currency if he can persuade libraries that all they really need is information of the sort and quantity which appear in the cataloguing-in-publication data. That plainly is available at the time the book is published.

I don't think he's yet seen that if that is so, there's no need for a BNB at all, because every library buying a book will have the cataloguing there. But if such a notion were to be both true and accepted, then the kind of cost comparisons which have led to the widespread acceptance of OCLC might prove to be the wrong ones, because it might still be cheaper for libraries not to look up OCLC records, but simply to copy the CIP data, or the Library of Congress data from the books which they have before them, and use lower paid, less trained staff to do it.

There may be other indicators in this direction. One of the aims of designers of large central computer systems has always been the minimisation of storage, in spite of the fact, clearly demonstrable at every stage over the past fifteen years, that the unit cost of mass storage in computers is falling, and will continue to fall in real terms, to the point now where I can buy a single hard disc to be driven by a micro-processor costing less than 100,000 dollars which is capable of holding a million Bodleian catalogue records – of course, we already have short records.

This may point to a division of function of our present catalogues: in-house short records, finding-lists, and bibliographical databases available on such commercial services as Lockheed, just like the abstracting services. You pay more for that.

There are other pointers, still, to different ways of using electronic hardware. There is now available a free-text searching engine, using at a key point decidedly *non-Von Neumann* logic.²⁴ And into this, you can

²⁴ This refers to the Memex Search Engine produced by Memex Information Engines Ltd. Jolliffe had just heard (in January 1983) of its existence and obtained information which was to lead to a Bodleian-based British Library R&D evaluation project (SI/G/627): the report by Geoffrey Neate and Lou

simply type catalogue records, the whole text of books if you wish, and you can search on the strings. You will be able to type in a MARC record, tags and all, and search on it as efficiently as typing in just the data with no tags, because what you would be generally speaking looking for is the information and not the structure.

There now exist browser terminals without keyboards – people touch the screen – where users can be guided and can guide themselves though catalogue searches. And there is much talk on the margins of that branch of study known as artificial intelligence of expert systems in which the expertise of a practitioner in a field such as medical or geological diagnosis is put into question and answer form.

It seems to me only a matter of time before someone tries to replace the reference desk personnel by their own encapsulated expertise, and I shudder to think where the place of librarians will be then.

[applause]

Burnard was completed after Jolliffe's death and was not published. Memex was based on piping pre-tokenised compressed text as fast as hardware would allow through the multiple parallel channels on a proprietary hardware board. The hardware was loaded with the user's query tokens and returned data addresses whenever these were detected in the compressed data stream. Memex was used by news and legal databases and by a telephone directory inquiries service.