The Shape of Libraries to Come

Dr R. Alston, British Library, Interviewed by A. Baron

The following is based on interviews with Dr Alston in 1991 and 1992. It was originally published as part of a longer report (*The Shape of Libraries to Come*, by Alexander Baron, InfoText Manuscripts, London, 1993, 1-871-47331-4). It deals with the problems of book classification, "information overload", database access and other subjects. The interview was conducted at a time when Dr Alston, now Director of the School of Library, Archive and Information Studies at University College London, had been conducting online "surgeries" with readers at the British Library for just over two years. These sessions continue and are held on two mornings each week. The range of subjects between early 1990 and late 1993 was extraordinarily wide-ranging; a few sample topics picked at random from the list are as follows: the liturgical texts and nineteenth-century British scholarship; magnetism and sympathy in European literature, 1700-1850; comparative linguistics – English and Russian; George Puttenham (1589) and Renaissance theories of rhetoric and stylistics; town planning in Hong Kong; personal accounts of survivors of German concentration camps in Poland; the growth of the sugar economy in the West Indies up to the abolition of slavery; fictional accounts of the "Mutiny on the Bounty". These represent only a small fraction of the range of topics covered. In his inaugural lecture at University College London in February 1993 ("The Battle of the Books"), Dr Alston elaborated some of the reasons why he believes that readers now, more than ever, need expert assistance with the sheer volume of electronic information available.

Dr Alston: On the whole, the more interesting kinds of topics that have been cropping up have been the ones where it has been extremely difficult to be of assistance using computers. One reader was researching the processes a script writer goes through when he adapts [sic] a play or a novel for a film. There have been some studies which come fairly close; I did actually find him two studies which were germane and which he didn't know about [on OCLC], but of course, one of the problems is that until a scholar defines a subject as being valid for research and enquiry, nobody uses that particular reference as a subject heading. So [in this case] we searched in OCLC under pictures, films, cinema and all the rest of it, and that was just no help at all. I find these the most interesting kinds of searches because I really don't enjoy doing basic elementary searches for PhD students; it's much more challenging and interesting to find someone who's actually working on something quite new which requires you to do some fairly sophisticated navigation using the databases.

Question: There was a review of the costs of offering this service in September [1990], I gather; what was decided as regards long-term policy, costs and otherwise?

Dr Alston: There is a serious problem here because if you charge the customer for the very same service which at the moment I'm providing free, you immediately change the nature of the transaction: instead of the scholar coming in, being somewhat laid back, perhaps pleased with whatever you do and all the rest of it, the moment money's involved and he has to pay, he's of course going to expect a professional response. You offer a service for which there's a fee, then you obviously have to satisfy the customer that what you're providing him with is worth what you're asking him to pay. Citizen's rights and all that.

That's fine in the case of myself doing a search, but I can't spend the rest of my life doing searches for readers in the Reading Room, which means training a whole group of curators who can take it over. And as I pointed out to the library, to train a whole group of curators costs money. You can't just take anyone in the library, put them in front of that terminal and tell them to get on with it because you'd get zero results, and the customers are not going to pay for zero results. So at the moment it continues to be free. I have to use discretion: Here a scholar comes in and says he wants every single book written about Samuel Johnson, and I can determine that that's 260, he can't have it free I'm afraid. We do exploratory searches for nothing, then if it's a very positive result and it's a large number, then he has to pay. So far I haven't actually found anyone unwilling to do that.

You can get onto anything by way of a database using JANET as a vehicle, and the proposals are, in the next couple of years, to extend the use of JANET beyond the academic network into the public sector and the private sector, which will then make it virtually a national network; so you would be able to obtain access to Cambridge University Library if [say] you were an ordinary citizen of Glasgow and you went into a Glasgow public library.

Question: Do you foresee a gradual merger of databases, as with police databases, for example? Do you think there will eventually just be one enormous database which anyone can log onto?

Dr Alston: I think, for a whole variety of reasons, that the drift will be in the other direction. There comes a point with the volume of data, where the sheer size is counterproductive to searching. You just get too much noise. If you were to collapse all the DIALOG databases that deal with the humanities into EPIC [OCLC] and then did a search on Samuel Johnson you'd get such a ridiculous figure that it would be useless. No one's going to sit and look at 43,000 entries on the screen: every article published and loaded on a database in the past 50 years; every journal of literary studies; every book. It becomes counter-productive; so I almost see there coming a point where even a database like OCLC would decide to split itself into groups: [for example] all books printed before 1850; all books printed between 1850 and 1939; all books published since 1939…simply to improve the efficiency of the search and make it easier to handle.

Question: Can you see any way in which the quality of online references can be sifted? For example, if you were to take the Prince of Wales as your subject term, you'd get literally thousands upon thousands of references. Is there, or can you see any way of sifting the references so that you can retrieve only the quality ones?

Dr Alston: Now you're into the subjective area; who's going to decide what is important and what isn't? If you're after scurrilous activities, proclivities, and you've already selected out all those references, then you'll get few results. I'm a great believer in extending the bibliographical record (which as you know is just a structured description of a book) to include something which you might call a summary, mini-review: a single paragraph, because what you cannot tell when you look at the record of a book on a database is – you can tell how many pages it has – but what's the book like? Is it trivial, is it well illustrated or anything of that sort: you've no way of knowing. And if that book is located in an American library 6,000 miles away you're not going to put through an inter-library loan just to have a look and then say five minutes later: "It's rubbish; send it back!". It could be a mini-review; it could even extend – and this could be automatic – an image of the table of contents. [If according to the title a book is about] *The Age of Johnson:* that extends throughout the last half of the eighteenth century: what's the book about? Is it social life, and if it's social life what aspects of social life? Is there a chapter on cooking? Is there a chapter on etiquette? Is there a chapter on sex? You don't know. Unless that kind of information is put into the record to give it added value, we're going to end up with huge databases with bibliographic descriptions which are not very helpful.

In the days when, with manual bibliographical catalogues, you turned to Samuel Johnson [for example], saw you had three pages to read, and you could think: "Ah, that's okay, I can manage that. That looks interesting…" then you go and check them. Nowadays on these big databases you're not talking about three pages, you're talking about 30 pages of references.

I think it's high time that we started introducing that element into the description otherwise we're going to be completely swamped with unnecessary inter-library loans. The real problem is that 20 years ago you could come into the British Museum Library (as it then was) fairly confident that if anybody had written a book about Anne Radcliffe (like our enquiry this morning) the British Library would have it. In other words it was a one-stop library. As long as you could find the books, you didn't have to go any further.

That is now finished; there's no one library on earth that can stock everything. Even great research libraries like BL are having to cut back on books printed in the Commonwealth, books printed in America…so the only way to find out what has been written about Anne Radcliffe or Samuel Johnson is to use a database. But if the database tells you: "This book is in Illinois; this book is in New York; this book is in Edinburgh…" – what are you going to do: request them? Only to open the package and say: "That isn't what I wanted". This is a waste of resources.

Question: There are developments in the online field though which may rule this scenario out: the *Bible* is already online. Now, there is no reason (in theory) why every book or periodical which is published by any major publisher should not be available on disk, and eventually online. It's got to come. So you could do a search in the library and if you like the look of a book you'll call up the entire text. Eventually we will reach the stage where it won't even be necessary to visit the library; you'll get up in the morning, turn on your bedside computer and link up to a database and read any publication in the world: book, dissertation, newspaper, magazine, in English, French or Russian. How far are we from that scenario in real terms?

Dr Alston: Technically speaking we're there; the problem will be having the equipment necessary to facilitate it. We're now talking about massive storage, storage that dwarfs Ministry of Defence computers. And this would create a very interesting situation, because if books were available, the whole texts in electronic form, the publishing business would disappear except for things like *Whitaker's Almanack*, railway timetables, *What's On in London* and all the things that you want to be able to carry around with you. Scholarly academic books would only need to exist in one copy. And if you did the economics of this you would then be told that in order to load the latest edition of Cambridge University Press's *Life of Samuel Johnson* online in an edition of one copy was going to cost [perhaps] £35,000. Now, who's going to pay that? You could say then that we would have to set up something like the Performing Rights Society and that the copyright of that book vested in the publisher or the author, and anybody who accessed the text of that book online would have to pay a fee.

At the moment we have no system that could conceivably cope with that. But after all, the Performing Rights Society exists, and they have a system which copes: you want to put on a play, you want to play a pop song – you have to pay your money. You want to read somebody's *Lovelife of Johnson*, it's going to cost you. That would be very interesting: again, how do you decide to cost access?

There would be huge savings: you wouldn't have to cut down half the world's forests in order to print 500 copies of books like the Cambridge University Press's *Life of Johnson*. Yes, there would be enormous conservation benefits – if we could do it.

Question: So really the only obstacles to be overcome are economic not technical ones?

Dr Alston: Well, there are two kinds of use of a book online: I want to look at page 6 – that'll cost you five pence. I want to look at pages 6 to 20 – that's going to cost you 14 times five? I want a printout of the whole book, I can't live without it – then you're going to have to pay for it…pay what? Exactly the same as you would if it had been printed in a limited edition of which one copy would cost £60? I don't know; nobody knows the answers to this because we don't have any precedent, we've no experience to build on.

Question: How has bibliography changed since the arrival of OCLC and other online databases? Is your job getting easier?

Dr Alston: It's a bit of yes and a bit of no. It's getting easier in the sense that you no longer have to trudge around the libraries of Europe and the world trying to find out what existed. I did this in my early academic period. That was great fun. The period between 1960 and 1970 when I must have visited and worked in nearly 1,000 libraries in the world from Japan, Australia, Europe, America, Canada: those were the fun days because you weren't just pressing a button on the computer and getting a listing. You were like the explorers, charting new territory. And that side of bibliographical scholarship has inevitably gone. No one can afford it anymore. Travel's too expensive and there are so many encumbrances now just getting across the Channel. So it's a trade-off: on the one hand, a lot of the drudgery has gone, on the other hand, a lot of the fun has gone, too. It's no fun sitting behind a computer terminal tapping at a few keys.

Question: Has the nature of bibliography changed as regards subject and category terms or is it all still *Dewey*?

Dr Alston: This is the most interesting thing about all this; because of the computer's ability to index different fields, we're no longer stuck with the tyranny of an alphabetical catalogue which runs from A to Z. We can have it that way if we want, but we can have it much more importantly by subject. But, there is no agreement (and I don't see how there ever can be any agreement), on a permanent way of describing the world's stock of books, any more than you can describe the world's stock of chairs. This chair's different from this chair: within a mile of this place you could find over 100 different varieties of chairs. We may all agree that they are called "chairs" because they are physical objects, but what about a book, which is a product of the human mind?

How are you ever going to agree on what you call André Malraux's famous book on culture, *The Voices of Silence*, a massive treatise on the whole concept of modern art; but what is it? You can't put it under art, you can't put it under culture, you can't put it under civilization; or aesthetics. What can you put it under? We have no word to describe that sort of book. There have been many attempts to invent certain terms to describe certain books, but none of them has succeeded. If you're too sophisticated with a search term (as OCLC sometimes is), then you'll never find the book because you can never predict how refined the person who catalogued it has been. If on the other hand you're too general, you get swamped. If you used the search term *God* for every book on every conceivable religion you'd get three million items; do you want to look at them all? It's totally ridiculous.

Question: That brings us conveniently round to throwing out the garbage. When I interviewed Buck Bloombecker for *New Computer Express* he told me that he keeps his office looking like a garbage tip, so that if anyone were to break in they'd think there was nothing of value there. But this is a serious problem now with regard to what is sometimes referred to as the information revolution, but which I call the information explosion. I have a catalogue here of a publisher which specializes in computer books. Microcomputers are a relatively new subject, it's only in the last ten years or so that they have ceased

to be a mystery for the man in the street. And not much more than 20 years ago they didn't even exist. Yet this catalogue contains more than a full page of listings for books on MS-DOS alone; that's an A4 page of small print. Even allowing for a lot of this material being duplicated, there is still an enormous amount of DOS here.

Dr Alston: Not as many as you've got cookbooks. Everybody writes a cookbook. Every part of the world you go to, the smallest village you can find, go into the local book shop and there is a locally compiled guide to the cooking of…(wherever).

Question: If anything this reinforces my point. There's so much information available. Walk into any large branch of W.H. Smith and as far as the eye can see you've got magazines: you've got two or three on astrology; you've even got two on astronomy, which is hardly a popular subject. You've got magazines for property owners abroad. If you were to put all that online, where would it all stop?

Dr Alston: When all this started we heard all this talk about the paperless office. Contrary to that prediction, the more computers are used, the more paper proliferates.

Question: It's not just paper though, it's this fascination with collecting information and the simple fact that there is more and more of it: the information explosion! It's nice to have this information, but what do you do with it all?

Dr Alston: The trouble is we are going to suffer, and in fact already are suffering, from information overload. This can impact on history, on what happens in the world. Gandhi gets assassinated on Tuesday night at a quarter to seven; it's on every television screen in the world five minutes later. In the nineteenth century news of the assassination of Gandhi would have taken three months to reach England. Now we have the theory of instantaneous access to information, which is what really worries me because the human being is not conditioned to respond to too much instantaneous information.

Question: Not only that, but, for example, bibliography is a highly specialized and very narrow field of human knowledge, but the point is, will it be possible in the future for anyone to be an expert on anything?

Dr Alston: That depends on how you define the word "expert". If you take any subject that is recognizable as a subject, you can find references to it online. Take a really abstruse scientific subject, say endocrinology, or even narrower than that, the endocrine system with special reference to the pineal gland, you're still dealing with an enormous volume of material, from Russia, Japan, India, America…on that tiny, minute subject.

Question: Which brings us conveniently to the question of the accuracy of information, both online and in the field of human knowledge in general. How much disinformation is there on databases and in books?

Dr Alston: An enormous amount. Subject classification is absolutely riddled with disinformation, even something as simple as authors' names.

Question: It all comes back to throwing out the garbage?

Dr Alston: Yes, but how do you do it? If you throw out the bath water you're liable to throw out the baby with it because in a sense, just as the experience of reading a book is a very private one, describing a book is equally private. We're heading for what the proponents of chaos theory have been telling us. Chaos! There's too much information; we can't possibly sift it or analyse it, and it's increasing, not decreasing. The more instantaneous communications become... the more the overload.

Question: The catalogue in the British Library is still only available online from 1975 onwards; when are readers going to get what you've got?

Dr Alston: The question is, supposing the library were to put ten carbon copies of what's in my office in the Round Reading Room, what do you think would happen? It would be tied up with people trying to work out how to get any information out of it. The biggest problem is getting everything the British Library has described in a format which can be mounted in machine-readable form. That's the problem. This is what an American librarian referred to as "the 98 per cent principle" which means that any library's catalogue only reveals 2 per cent of what's in that library.

Question: Something like 60,000 books a year are published in this country alone. Bearing in mind that this includes such abstruse titles as *European Concrete Directory*, the *International Directory of Crematoria* and *Who's Who in Poland*; and that such books as the latter are not necessarily published every year, who reads them, who commissions them? They are so extremely specialized that their print runs must be infinitesimal. How do the economics work? What sort of print run will a book like the *International Directory of Crematoria* have, or *Who's Who in Poland?* Who commissions these books? How are they published? How does the publisher make a profit?

Dr Alston: What publishing has always sought to do, right from the beginning, is to identify a market. Every publisher has to ask the question: If I publish this book, who's going to buy it? And if he can think of absolutely nobody who will buy a certain book, then there's not much point in him publishing it. So, if books appear on a subject as arcane as you've suggested, the chances are that that publisher thought he had a market for it. If you think of the population of the English-speaking world, which is now 500+ million, there's room in that 500 million for books on absolutely anything. It's your magazine problem again. They only publish these magazines because there is a community out there interested in the subject in question. And if there are more than 500 people out there interested in it, then there's room for a journal. How you get at those people is another story, because W.H. Smith only stock a fraction of what actually exists. If you went into parts of London with Asian communities, Islamic communities, or whatever, you would find a whole range of other books and magazines on weird and wonderful subjects you'd never heard of.

Question: Most non-fiction books are commissioned either from academics like yourself or from people knowledgeable in their field.

Dr Alston: Yes. I happen to know somebody who is in medical publishing, and they publish huge, expensive books with colour illustrations on every aspect of medicine. Every one of them is commissioned. The publisher knows that there is a market for a text on [say] dermatology which will be bought by students of dermatology and medical libraries. They do their market research very carefully – how many medical libraries are there which will buy the latest, definitive book on dermatology? Three hundred? – How many students of advanced study of dermatology in the English-speaking world? Four hundred? – Edition, seven hundred?

There's an enormous funding of academic research books by the funding councils like the Social and Economic Research Council and so on. These tend to be books on subjects where the print run is so small (although significant) that they could never be published commercially. And that's a good thing.

Question: Otherwise they'd never get into print?

Dr Alston: No. But having said that, there is a tremendous urge now on the part of writers who apply to these research foundations to think of a subject on which nothing has been written.

Question: Is it worth it?

Dr Alston: If you take the long view of history, maybe; if you take the short view – we're reaching the point now where if we don't stop polluting the atmosphere and tearing down the rain forests, all this is going to be to no avail anyway, because we'll soon be living on an uninhabitable planet.