

Key factors in the development of digital libraries

PROF. JOHN MACKENZIE OWEN¹

Abstract

The library traditionally has performed a role within the information chain, where publishers and libraries act as clearing houses between authors and users. In this model various institutional parties perform specialized functions. Publishers are especially oriented towards authors, for whom they perform a dissemination role, manage quality control, create the 'canonical archive' and offer authors recognition through publication of their ideas and findings. Libraries are especially oriented towards the end user, for whom they provide selective filtering of information, storage of information resources, a variety of services and support. The interrelationships between authors, publishers, libraries and users have resulted in a highly efficient logistical system for the distribution of knowledge in printed form.

This system now seems to be changing under the influence of the transformation from print to digital information. In this paper the main issues underlying this transformation, and the impact of these on the future of libraries, are examined.

The key technological driving force is the combination of digital information and the digital network as a distribution medium. Because both the media and the distribution channel are becoming digital, we now are moving towards a digital information chain where traditional institutions such as publishers and libraries will have to redefine their roles. Within the broader context of digitization we can identify a number of key developments which libraries should take into account. These include: interactive media, new applications of information technology,

¹Department of Information Science, University of Amsterdam, e-mail: mackenzie.owen@kub.nl

new co-operative organizational models, knowledge management as a new context for the library function, and the emerging cyber generation. These developments are discussed in more detail in the paper.

The paper concludes with some thoughts about the library of the future. The future library can be described by the words 'digital', 'virtual' and 'distributed'. Its most striking characteristic will be that it will not perform a storage function, i.e. there will be no digital collection. Rather, the library will act as a contract manager, managing license agreements with copyright holders which allow the library's patrons access to commercial information resources.

Finally, the major challenge for libraries will not just be to move from traditional models to the digital, virtual, distributed model. The real challenge will be to do this in a way which offers added value to the user.

Libraries in the information chain

The library traditionally has performed a role within the information chain, where publishers and libraries act as clearing houses between authors and users. In this model various institutional parties perform specialized functions. Publishers are especially oriented towards authors, for whom they perform, in the case of science, a dissemination role, manage quality control, produce the 'canonical archive' (i.e. the global archive of scientific knowledge) and offer authors recognition through publication of their ideas and findings (Rowland 1997). Libraries are especially oriented towards the end user, for whom they provide selective filtering of information, storage of information resources, a variety of services (such as cataloguing and indexing, document delivery etc.) and support. The interrelationships between authors, publishers, libraries and users (including other parties as well, such as bookshops and subscription agents) have resulted in a highly efficient logistical system for the distribution of knowledge in printed form.

This system now seems to be changing under the influence of the transformation from print to digital information. In this paper I shall examine the main issues underlying this transformation, and the impact of these on the future of libraries.

Digital information and networks

First, then, what are the main issues underlying the transformation from print to digital information? The key technological driving force is the combination of digital information as such and the network as a distribution medium. It is now quite clear that more and more information which used to be printed on paper, is now being published in digital form. Currently most publishers, especially the main scientific publishers, prefer a dual mode: journals are published both in print and in digital form. In many cases, the digital version is derived from the printed form (e.g. through scanning), and at least there is a tendency to ensure that the digital version is an exact copy of the printed journal. It is highly likely that within a number of years most major scientific and professional journals will be available primarily in digital form. By that time, the formats of these journals will move ahead from the characteristics of print, and begin to make more use of the possibilities offered by the digital form, e.g. flexible publishing schedules, interactivity, multimedia etc.

The off-line distribution of digital information, e.g. via cdrom, only seems to have a future as a distribution format between systems and for specialized information products, e.g. products of which small numbers of users make heavy usage. In general, it is the network, not off-line media, which holds the future as a distribution medium. We are therefore now moving towards a digital information chain where traditional institutions such as publishers and libraries will have to redefine their roles.

The network now means: the Internet. In future we shall see a larger number of networks, trading off speed and quality against price: fast advanced networks for power users in industry and universities, cheaper but slower and less reliable networks (such as the current Internet) for household use. That will not necessarily lead to fragmentation of the available information. Any information site could be accessed through any of the networks; the only difference will be quality and price of the delivery mechanism, not of content.

The significance of digitizing and networks is well known: global accessibility. Any information, wherever it is located, can be accessed from any other location, usually at an acceptably high speed. One consequence of this is information overload: how to choose if everything is available? Here the library can contribute significantly by offering advanced search and selection mechanisms, tailored to individual or group information

needs. Another consequence is of also significance to libraries. Traditionally the library acts as a store or memory, ensuring that important information is available locally, as near as possible to the user. In the digital world location and speed of delivery are not an issue any more. Libraries will therefore have to ask themselves whether their storage function needs to be maintained, and more specifically whether it makes sense to develop new storage facilities for digital information. That is an issue I shall come back to.

Key developments

The Internet has stimulated many developments which influence the distribution of information through the digital information chain. I would like to mention five of these:

Interactive multimedia

Although we have had multimedia (the integration of text, sound, still images, moving images and embedded software applications) for at least a decade, the real significance of multimedia comes into the open through the network. The result is *interactive multimedia*: a far richer variety of media types and media uses (e.g. networked radio and TV, music clips, games etc.), and the use of multimedia in the communication process as a component of both human-machine and human-to-human interaction. We can now add a comment to an article, send a media-rich reaction or a simulation to the author, speak to him or here, even look at each other, all through the network. Dissemination through the information chain is no longer confined to printing words on paper, but has become a media-rich interactive process.

This has a bearing on the nature of content (cf. the discussion between Zariski, 1997 and Hibbits, 1997) and on archiving issues, and by extension on the role of publishers and libraries. The combination of networks and multimedia leads to a far more expressive and therefore personal mode of communication than the traditional print information chain can offer.

Information technology applications

New applications of IT are developing which enable interactive, multimedia communication via networks. The most significant of these is the

browser. The browser is now becoming the key end-user application. The modern browser application serves as a local and external file directory, allowing users to access and manage information files. The browser provides hypertext links allowing an associative walk-through in information space. It offers the user access to search machines, it can be used to obtain documents, to view and to print them. It integrates a message system allowing one to send and receive multimedia e-mail. It acts as a document creator, employing the future global standard document format (i.e. HTML and its successor XML). It incorporates all the functions a user can perceive: navigating, browsing, viewing, filing, creating and communicating. It can also serve as the interface to more specialized applications, ranging from online catalogues and databases to administrative applications, office functions (such as word-processing and spreadsheets), simulation software etc. In fact, a browser embedded in an underlying operating system is all the technology you need at the level of a personal workstation linked to the network, as both Sun and Microsoft seem to understand extremely well.

Other applications of IT lead to search engines, information filtering, knowledge extraction and intelligent agents which perform user-related tasks automatically (one of the finest visions on agents is given by Negroponte, 1995). Finally, IT-developments offer virtual integration of geographically distributed systems, e.g. through the Z39.50 protocol. Through this mechanism, the user need not perceive the world as consisting of different organizations and services. Rather, users can create a personal service based on an individual combination (or 'coalition') of service providers, search engines, and intelligent agents, available as a coherent, integrated service through the browser on their personal computers.

Co-operation

Co-operation is another issue underlying the development of networked digital information services. It is long ago since the local library was the single source of information available to an individual user, and since information was not available if it was not available in the library. In the networked world all information can be available to everybody. Information technology, digitizing and networks lead inevitably to co-operative relationships between institutions within the information chain. In as far as users really will need library services in future, they will

not be content with the service of a single library. Hence the increasing co-operation between libraries, as exemplified in the many co-operative digital library projects carried out within the framework of R&D programs (e.g. the European Libraries Program and the E-lib program in the UK, as well as many other national programs).

In addition to 'horizontal' co-operation between libraries (and, although as yet far less so, between publishers), there is increasing 'vertical' co-operation between institutional parties in the information chain. The prime example of this is the tightening relationship and increasing interdependence between publishers and libraries. There are many aspects to this. For publishers to succeed in moving users over to digital publications, they need the support of libraries, who for many reasons (especially administrative and financial) continue to provide the access point to published information for the end user. This means that publishers depend on the deployment of IT by libraries, the development of digital libraries, and the persuasive force of libraries with respect to users. It also leads to new contractual relationships, such as licensing as an alternative to the traditional subscription scheme, which entail a much more detailed regulation of rights and duties between the publisher and the library. And finally, libraries and publishers are competing for the same functions within the information chain an issue I shall return to later on with respect to the storage function.

Another area of co-operation which deserves mentioning is that between libraries and computer centres. Libraries need the information infrastructure provided by computer centres in order to make their services available to the user. Computer centres need libraries as content providers in order to justify their investments in networks, personal computers etc. Supporting information services through an IT-infrastructure, rather than traditional computing and data processing, is rapidly becoming the primary function of the computing department. Co-operation is in the interest of both parties. Experience from our consulting practice shows that both perform most successfully when they work together.

Knowledge management

The communicative aspects of IT and the drive towards co-operation have a very specific effect on libraries, namely that their boundaries are becoming far less sharply defined. What is the library when its services

are offered by a multitude of organizations, when its holdings are scattered over the globe, when it appears to be no more (or less) than a function on the desktop computer, called up by clicking on an icon with the label 'library'? How can an outsider distinguish between what is done by the library and what by the computing department when looking at the screen? Is the library still a library when it does little more than pay the bills for materials accessed by its users on the computer of a publisher? Traditional concepts of the library appear to be no longer valid in the digital world. This is all the more true when we look at the new concept of knowledge management.

Knowledge management is an approach based on the central role of knowledge in organizations, with the objective to manage and support knowledge work and to maximize the added value of knowledge for the organization. Knowledge management aims at identifying and analysing knowledge and knowledge work, and at developing procedures and systems for generating, storing, distributing and using knowledge in the organization. Knowledge management is predominantly a new way of thinking about information in organizations. It focuses on the strategic importance to the organization of the quality of knowledge, its availability and use.

One of the main carriers and distributors of knowledge within an organization is the human being. Knowledge is contained 'between the ears' and is transferred from one person to another in a variety of ways. One aspect of knowledge management is therefore 'human resource management'¹, ensuring that people within the organization have the right knowledge and that they communicate their knowledge effectively. The other carrier of knowledge is, of course, information. The advantage of information as a carrier of knowledge is that it is explicit, it can be stored, searched for, retrieved and processed. That is precisely what libraries traditionally have been doing, and they seem right in saying that they are and always have been 'knowledge managers'.

But that is not entirely true in the context of our current understanding of knowledge management. Knowledge management is an integrated approach, encompassing both 'human' and 'documented' knowledge, both internal and external knowledge, covering content, behaviour and support systems. Knowledge management is primarily the responsibility of the top management level within an organization. Therefore, librari-

ans can play a role in knowledge management as managers of 'knowledge information'. But they have to understand that they then act as a component of an integrated knowledge management process, working together with others (e.g. in areas such as human resources and information systems) under control of higher management levels. They also have to understand that 'knowledge information' includes a far greater variety of information resources than traditionally found within a library, many of which are to be found in other areas of the organization (e.g. internal information on processes, products, projects and people).

So to summarize this issue, knowledge management is yet another example of libraries performing a valuable function for end users rather than being identified with a precisely defined organizational entity. Library functions remain extremely useful and can be expanded to areas not found in the traditional library. But the library as an autonomous institution seems to be dissolving into a co-operative structure of 'knowledge process supporters'.

The cyber generation

In looking at developments and their impact on the information chain, one naturally tends to look at information technology as a driving force with a direct bearing on information institutions. However, there are also important indirect forces which have to be taken into account. The most important of these is the impact of IT on the people who constitute the human factor in the information chain: authors and end-users. The main difference between the two is that new generations become end users before they become authors, if at all.

Children are now becoming used to IT ? personal computers, computer games, the Internet ? at an increasingly young age. They regard these things as a source of information, as a medium for communication and as a tool for learning. Most students entering the universities are now highly proficient in using a computer, and often have more experience with the Internet than their tutors. Digital information and interactive multimedia are perfectly normal things to them. As the computer and the Internet become household goods and primary and secondary education joins the bandwagon of the Internet and computer-assisted instruction, digital information will be more normal for future generations than printed information. We will no longer need to teach them how to use the

computer or browse the Internet; we may have to teach them how to browse a journal, how to look up something in a book, perhaps even how to read the printed word.

The future cybergeneration will certainly not accept the traditional, print-based library as the single (or even as an important) source of information. If we cannot offer them digital libraries and well organized access to networked resources, they indeed will begin to ask us questions. These aspects, amongst many others, are the starting point from which to discuss the future of the library.

The library of the future

What, then, will the library of the future look like? In general terms, the future library will have three main characteristics, which can be summarized as digital, virtual and distributed.

The library will be *digital*, i.e. it mainly will be based on digital information resources. This is due to the inevitable move from print to digital publishing. The print collection will therefore gradually become less important. In view of the significant differences between the print and digital medium (in terms of acquisition, handling, storage and delivery), one can also expect an organizational division between the digital library and archival, print-based libraries.

The digital library will be *virtual*: it will be available on the user's desktop through the network infrastructure. 'Going to the library' will not mean going to a 'physical' library location, but accessing a function, 'clicking on an icon' on the computer screen. The 'real', physical library will be invisible to the end user, except in its archival role as a collection of print publications from the pre-digital era.

Finally, the digital, virtual library will be *distributed*: no single library can or should wish to offer a full range of services catering for all the information needs of its users. Instead we shall see coalitions of libraries, often within a specific subject domain, offering co-operative services in a transparent way, i.e. without the end user ever noticing that the various resources and services offered come from different locations and different organizations.

Creating the digital, virtual, distributed library (or *networked library* as I prefer to call it), is a major challenge to libraries of the present. In doing this, various issues have to be taken into account.

Perhaps the most significant issue is that the traditional storage function of libraries, the *collection* - will disappear. Storing print items locally in the library makes sense, since availability of print information is directly related to the physical distance to information items. The drawback is, of course, that items are duplicated in many locations, and that storage costs have to be carried by the individual library. In the digital, networked era physical distance is of no importance. Any document anywhere on the network can be accessed from any location. Duplication of storage now becomes a matter of network efficiency, and is no longer required at the user's library, except perhaps for very heavily used materials, and only for as long as usage exceeds a certain level. From an organizational viewpoint, the best location for storage and access of publications is with the originator, i.e. the publisher. This is precisely what is happening with the large scientific publishers such as Elsevier (Science Direct) and Springer Verlag (Link). An important side-effect is that the whole area of cataloguing (or *metadata* as we call it nowadays) will also move from the library world and even authors.

The shift from libraries to publishers as storage collections leads to another significant change for libraries. Instead of having a role in creating and maintaining collections, libraries will act as *contract managers*, managing license agreements with publishers. License agreements are rapidly becoming a substitute for subscriptions. Subscriptions are agreements to purchase materials (e.g. journals) under certain financial conditions. Licenses are contracts which regulate access and usage rights for end users. They specify who are allowed to access which materials, and what users are allowed to do with them. The advantages of such contracts are that the publisher need not deal with individual users, and that users can access a wide range of materials through a single access point, using a single identification and without having to pay for each individual transaction. It should be noted that a license can cover both unlimited access (for a specified user group), or pay-per-view (payment based on the volume of access). In this role, the library can be regarded as a *clearing house* for access rights to networked information.

A further issue for libraries is related to the expanding range of digital information resources. The range of digital resource types is far richer than in the world of print, where one primarily has to deal with fairly well specified formats such as books and journals. Digital information resources include such diverse things as electronic journals, websites, pre-

print archives, individual multimedia objects, software, databases, all kinds of services, etc. In addition, there is a wide range of technical formats, and the underlying technology is changing rapidly. The library therefore has to manage a dynamic technical infrastructure in order to maintain access to information resources for its users. This also requires new types of systems within the library, including search engines, intelligent agents, information filters and catalogue systems which can cope with the increasing diversity of digital resources.

The role of the digital, virtual library within organisations - e.g. the 'corporate library' - will have to be defined in terms of knowledge management. Ultimately, the library will have to take on the role of 'content manager' of the organizational Intranet, ensuring that the wide range of knowledge resources required by the organization is available in an easily accessible and cost-effective way.

A final issue which has to be mentioned, is the changing information behaviour of users. I have already pointed to the oncoming 'cybergeneration'. But even existing users are changing their information behaviour under the influence of networks and digital media. Libraries will have to think hard about how to support these users, and how to offer services which offer added value above direct access to the network.

Conclusion

In this paper I have identified the underlying issues which libraries have to keep in mind when thinking about their future. I have also discussed a number of characteristics of the library of the future. The fundamental underlying issue is the shift from print to digital information, and the use of networks for the access to and distribution of information. Other issues include the increasing variety of information media, especially interactive multimedia, the role of IT-applications, the emerging concept of knowledge management, and the changing attitudes of users (especially of the younger 'cybergeneration') towards information and information technology.

The library of the future will be digital, virtual and distributed. The library will no longer be 'a place to go to for gathering information'¹, but have been transformed into a co-operative service available at the user's desktop. The storage function will disappear from libraries, to be taken over by publishers. Cataloguing will be done at the source, either by the

author or the publisher. An important new role for libraries will be in managing access rights, through licensing, for end users. Within organisations, libraries will play an important role as managers of the content available on the Intranet. Finally, the major challenge for libraries will not just be to move from traditional models to the digital, virtual, distributed model. The real challenge will be to do this in a way which offers added value to the user. Otherwise, the Internet itself will become the global library and libraries will disappear. I am convinced that would not be of benefit to the user.

References

Hibbitts, Bernard (1997) - E-Journals, Archives and Knowledge Networks: A Commentary on Archie Zariski's Defense of Electronic Law Journals. - In: First Monday, issue 2-7, 1997. [http://www.firstmon-day.dk/issues/issue2_7/hibbitts/index.html]

Negroponce, Nicholas (1995) - Being digital. - Vintage Books, 1995.

Rowland, Fytton (1997) ? Print journals: fit for the future? ? In: Ariadne, issue 7. [<http://www.ariadne.ac.uk/issue7/fytton/>].

Zariski, Archie (1997) - Never ending, still beginning: a defense of electronic law journals from the perspective of the E-law experience. - In: FirstMonday, issue2-6, 1997. [http://www.firstmonday.dk/issues/issue2_6/zariski/index.html]