Libri, 1999, vol. 49, pp. 181–191
Printed in Germany · All rights reserved

Libri
ISSN 0024-2667

Digital Libraries: Assumptions and Concepts

DAVID BAWDEN AND IAN ROWLANDS
Department of Information Science, City University London, UK

This paper clarifies some terms and concepts surrounding the idea of the 'digital library'. A set of twenty underlying assumptions, often expressed implicitly rather than explicitly in the literature, is outlined, and variant forms are compared. There follows an explication of a series of terms, essential to the understanding of the concept of the digital library and its components, which are often used nebulously or with different meanings. It is suggested that the lack of a common conceptual framework for the subject presents problems for the practitioner as well as the theorist.

Introduction

> ... what is a digital library? Conceptions differ. Approaches differ. Realizations differ. (Saracevic 1999)

The purpose of this paper is to clarify the meaning of the phrase 'digital library' and related terms, and to examine some common assumptions underlying the concepts. A second paper builds on this examination to propose and partially validate a high-level conceptual framework to aid in the better understanding of the idea of the 'digital library' (Rowlands and Bawden 1999). Both papers are based on a report commissioned by the British Library Research and Innovation Centre (Bawden and Rowlands 1999), but extend the scope of the material presented there, and include material not available when the initial report was prepared.

The digital library concept is of enormous social and economic significance. It promises to transform the way that services are delivered to the public (witness *New Library – The People's Network* and other recent public policy initiatives in the UK) and, potentially at least, it redefines the nature of the relationships between information users, providers and intermediaries. Not surprisingly, a considerable body of research has been carried out in the digital library domain. Contributions come from a wide range of disciplines, from sociology to information technology. For overviews, see the books by Lesk (1997) and Stern (1999), the proceedings edited by Aparac and others (1999), the articles by Kuny and Cleveland (1998), Chowdury and Chowdury (1999) and Borgman (1999), and the report by Brophy (1999). An examination of the published record suggests that the current digital library research effort is highly fragmented and characterised by a wide diversity of assumptions, definitions and views. Brophy (1999), in referring to the 'confused terminology' of the field, attributes it in part to the different perspectives from which commentators approach the subject. He identifies three main points of view: librarianship, within a digital environment; origins unrelated to librarianship, from computing research to e-business; and social drivers – policies, norms and cultural imperatives.

In this paper, we have tried to expose some of the assumptions – and sometimes, perhaps, misconceptions –which have attached themselves to the discourse of digital libraries, and to come to grips with this confusion of terminology.

Ian Rowlands, Department of Information Science, City University, Northampton Square, London EC1V 0HB, UK. Tel: +44 207 477 8389. Fax: +44 207 477 8584. Email ir@soi.city.ac.uk
David Bawden is Course Director for City University's MSc/Diploma course in Pharmaceutical Information Management, Department of Information Science, City University, Northampton Square, London EC1V 0HB, UK. Tel: +44 207 477 8390, Fax: +44 207 477 8584. Email: db@soi.city.ac.uk

Assumptions underlying the digital library concept

There are a number of commonly held assumptions underlying the writings in this area. Some of these were identified by Levy & Marshall (1995), who examined, in particular, three assumed characteristics of digital libraries:

- digital library collections contain fixed, permanent documents;
- digital libraries are based on digital technologies;
- digital libraries are to be used by individuals working alone.

They went on to show that each of these could not be justified as a universally applicable rule.

These may be contrasted with the three 'general principles' guiding the British Library's digital library programme (Lang 1998):

- the digital collection may be created and produced in a variety of different places, but will be accessible as if it were a single entity;
- the digital collection will be organised/categorised/indexed for easier access than is possible from its original point of production;
- the digital collection will be stored and maintained in such a way as to ensure that it will continue to be available long after the period of its immediate currency.

These typify statements of purpose for particular digital libraries: while they are not in themselves definitions, they imply a good deal about how such libraries are regarded by their creators.

We have identified a set of twenty assumptions, from the diverse literature of digital libraries. While they are not universally shared, as those quoted by Levy and Marshall were suggested to be, nor restricted to particular initiatives like the British Library's, they underlie all current writings, presumptions, and project planning in the area. Some of them have been disputed by individual commentators, some research projects have explicitly considered alternatives to them, and, in a few cases, diametrically opposite assumptions may be espoused. But, overall, these are the prevailing views among writers on the topic, and among the potential customers of such entities.

In the following section, the twenty assumptions are presented and briefly discussed. They are divided into four loose categories: basics; content; services and use; and research and evaluation.

Note that the use of the term *traditional library* refers to a pre-digital library, i.e. one based on a physical, largely printed, collection. This term is preferred to 'real library' sometimes used in the literature, since the latter conveys the implication that the digital library is in some sense an inferior, or incomplete, version.

Basics

1 – There is (or will be) such a thing as a digital library.

A 'digital library' may be understood in different ways, and may be named differently; Chowdury and Chowdury (1999) and Borgman (1999) draw attention to both complementarity and contradiction in various definitions. Terms such as electronic library, virtual library, hybrid library, gateway library, library of the future, and library without walls are used, sometimes synonymously with digital library, sometimes to denote a subset, or a superset, of it, sometimes to denote a rather different concept.

2a – Most, if not all, libraries will progress steadily to a 'fully digital' state over time.

2b – Most, if not all, libraries will remain indefinitely in a 'hybrid' state, with digital and non-digital resources offered side-by-side.

These conflicting assumptions are freely offered as alternatives in the literature. The first seems to be favoured, in general, by those working with scientific, technical and commercial material, especially those from a computing background, and those involved in 'leading edge' research; the second by those involved with research-based collections, especially in the humanities, by those from a librarianship background, and by those involved in service development. Espousing one or other of these assumptions is equivalent to a belief in the digital library and the hybrid library model respectively: Odlyzko (1997) and Crawford (1999) are good examples of the two viewpoints.

3 – Implementation of the digital library concept implies organisation and cultural change.

While this seems to be generally agreed, there is some difference in emphasis as to whether this

change affects primarily library / information professionals, users, or the whole environment.

4a – The digital library builds upon the traditional strengths of the library/information professions.

4b – The digital library is a catalyst for the development of the library/information professions in entirely new ways.

4c – The digital library marks the end for the traditional library/information professional.
Examples of all three of these assumptions may be found: the first most prominently in the writings of mainstream librarians, the last in those of the professions, particularly informatics, which claim to supplant them. The middle alternative is espoused by those who foresee the emergence of the cybrarian (a librarian in cyberspace), the hybrarian (librarian in a hybrid library), and similar species; see, for example, Hey and Wissenburg (1998), Cronin (1998), Garrod (1999) and Oppenheim and Smithson (1999).

Content

5 – Digital libraries will contain material in digital form.
Though obvious, this is, interpreted in different ways. Some commentators suggest that a digital library contain *only* digital material, current examples being merely transitional stages towards this (see, for example, Odlyzko 1997). More common is the assertion that print and other non-digital media will survive, and indeed flourish, alongside digital material because of their appropriateness in particular settings and for particular needs (see, for example, Gilster 1997 and Crawford 1999).

6 – A digital library will have an identifiable collection.
Although in this instance, the concept of 'collection' will be associated with matters of access, rather than physical location.

7 – Digital libraries will contain documents.
By documents, we mean discrete, identifiable, and relatively permanent, information artefacts. These documents will be varied in nature, but will certainly include items analogous to printed books and journals and possibly (certainly, according to some commentators) datasets of various kinds – cartographic, environmental, scientific, statistical, etc. It is worth noting that, as Brophy (1999) reminds us, Licklider's pioneering writings of 1965 referred to the handling, not of documents, but of 'the facts, concepts and ideas that lie behind the visible and tangible aspects of documents.']

8 – There will be a continued emphasis on text in digital libraries.
This will be true except, perhaps, for specific application areas, e.g. geographical information systems, or genome databases, where text will be a minor part.

9 – The contents of digital libraries will result from a publication process.
But 'publication' must be understood loosely, so as to allow for the modification, varying in degree according to the view of the commentator, from the traditional author-editor-publisher chain.

> Over the next five years, the digital library will move from concept to reality. As it does, the roles of many in the information chain will change; notably secondary publishers, subscription agents and document supply services (Smith 1998).

On the other hand, some digital library initiatives may be seen as automation of the *status quo:*

> Many digital library prototypes merely automate existing publishing practices or focus solely on the digitisation of publishing cycle output, not capturing elements of the input (Esler & Nelson 1998).

10 – Digital libraries will be concerned with preservation of material.
This will involve some aspects of the care of physical materials (e.g. optical disks) familiar to the traditional library, but also tasks such as refreshing and migrating digital materials, and possibly the preservation of IT equipment. There is also the new perspective presented by the idea of a collection defined by access, when the resources are not within the custody, in any sense, of the library.

Generally, this assumption is made by those from a library background. The computer science perspective on digital libraries lays greatest stress

on their dynamic nature, and is relatively unconcerned about long-term viability or resources.

Services and use

11 – Digital libraries will be used in much the same way as traditional libraries.
That is, they will be used. by individuals working largely in isolation and on essentially 'scholarly' tasks. This perspective has much to do with the fact that most work on digital libraries has been carried out in a higher education environment. Even here, studies in traditional libraries have shown that this is an over-simplification, and that much 'library work' is co-operative and problem-oriented (Levy & Marshall 1995). This assumption is challenged by some commentators who see the digital library as changing the place (or perceived place) of the library in the working lives of its users. Buttenfield (1999) describes as 'illogical' the assumption that user needs in the digital library mimic the needs in the physical library.

12 – A digital library will include an equivalent of the traditional catalogue.
It will require a richer set of metadata descriptors and standards for dealing with the variety of information artefacts to be identified.

13 – A digital library will provide a range of searching and browsing tools
A digital library, more than a traditional library, will be unusable without a full array of such tools.

14 – Digital libraries will provide integrated services, ideally approaching 'seamless' integration.
Usually this is taken as a given, though doubts are expressed as to the feasibility of total integration, and, even, to its desirability. The last point reflects concerns that as the seeking and viewing of information becomes standardised, and its presentation increasingly uniform, valuable cues for evaluation and analysis may become lost.

15 – Digital libraries will offer customised interfaces for particular users.
These will be in ways not possible with traditional resources. These assumptions have spawned a number of approaches, including the concepts of 'user profiles' and 'information landscapes'.

16 – Digital libraries will act as gateways or portals.
Although traditional libraries have always provided access to the wider information landscape for their users – most obviously by the provision of inter-library loans – digital libraries are able to do this in a fuller way. In principle, users need not be aware of when they are entering or leaving their own library's collections.

17 – Digital libraries will support the analysis and processing of information.
Digital libraries will be able to support analysis and evaluation of information as well as searching and browsing in a more direct way than a traditional library. This is made possible through the provision of appropriate software tools. They may be described as 'passive' or 'active', depending on the way they distribute electronic materials (Hayes 1997), but will certainly have the capability to provide a much more proactive supply of information, particularly by 'push' technologies (Pedley 1999), than traditional libraries. They may also, depending on their subject material, support tools for data mining and information discovery.

18 – Digital libraries will promote information literacy.
This will be achieved by playing a more active role than the traditional library in the promotion of critical thinking skills and of information literacy (Bawden 1999).

Research and evaluation

19 – It is possible to do research on digital libraries.
But this means different things to different researchers and funding agencies. In particular, two main forms of research can be identified in the literature and in project reports. One, generally based on advanced informatics, is deliberately innovative in seeking to set up new structures and environments, to prove new concepts, and to avoid automating the *status quo*. The second is pragmatic, concentrating on practical applications, and the piloting of new applications, and focusing on near-market technologies and cost-

effectiveness. Rusbridge (1998) contrasts the two approaches, exemplifying the former by the projects supported by the National Science Foundation and similar agencies, in the USA, and the latter by the UK e-Lib programme.

20 – It is possible to evaluate the effectiveness of digital libraries.

Some digital library research projects have had elaborate evaluation processes built into them from the outset. However, there is no consensus on the methods and metrics which are appropriate for evaluation in this area; 'traditional' measures of library patron satisfaction, and of information retrieval effectiveness, are clearly inadequate in a distributed digital library environment (Buttenfield 1999, Chowdury and Chowdury 1999).

It is evident that very few, if any, assumptions regarding digital libraries are shared, without qualification, by all writers on the subject. While this does not mean that there is a fatal flaw in the concept – rather that it is an area characterised by diversity and controversy – it does imply a need for care in understanding the implicit or explicit assumptions underlying writings on the topic.

Explication of terms and concepts

In this section, some major concepts associated with digital libraries are introduced and discussed. This is not a set of *approved* definitions. Rather, it is an attempt to explain the way in which these terms are used by various commentators to dispel the confusion caused by varied use.

This section is in two parts. The first deals with the digital library itself and with some other terms describing apparently equivalent or similar concepts. The second deals with some associated terminology and concepts.

The digital library and similar concepts

The use of varying terms over time in the literature was examined by bibliographic database searches. The association of ideas with each term resulting from analysis of Internet material is shown in Appendix 1.

The digital library is by no means a new idea

> Grand visions in 1960 led to the first development of text search from bibliographic databases to full text retrieval. Next, research prototypes catalysed the rise of document search from multimedia browsing across local area networks to distributed search on the internet. By 2010, the visions will be realised, with concept search enabling semantic retrieval across large collections (Schatz 1997).

Brophy (1999) traces the origins of the concept of organised collections of digital information to Licklider's influential writings of 1965. Crawford (1999) suggests that the peak period of belief in all-digital libraries came between the years 1990–1994. Since then the trend has been to think in terms of mixed, or complex, collections; the 'digital library in the classic sense', i.e. solely digital material, being used by only 'the most extreme proponents'. Be that as it may, the term is still widely used, though with a potentially confusing range of meanings.

In a 1995 statement, the US Association of Research Libraries noted that there are many definitions of a digital library, and that the terms 'virtual library' and 'electronic library' are often used synonymously. They identified five – very general – elements common to all definitions of the digital library current in the first half of the decade

- the digital library is not a single entity;
- the digital library requires technology to link the resources of many libraries;
- linkages between digital libraries and information services are transparent to users;
- universal access to digital libraries is a goal;
- digital library collections are not restricted to document surrogates but include digital artefacts that have no printed equivalent.

Some definitions of the digital library focus mainly on the technology

> A digital library enables users to interact effectively with information distributed across a network (Schatz 1997).

Others hint at wider contexts

> The digital library is a recent term used to refer to information systems and services that provide electronic documents from dynamic or archival repositories ... however, if digital libraries are narrowly defined, then we lose the ability to learn about key issues from previous research, theory and professional practice in information science and librarianship (Elliott & Kling 1997).
>
> The notion of the library as a space is better agreed upon than the notion of the library as an information repository (Ercegovac 1997).

Or stress that digital libraries are in fact better regarded as socio-technical systems

> ... rather than looking at the digital library as a specific technology, the present study takes an organisational approach, placing the digital library in the context of a specific learning and teaching environment (Travica 1997).

> [digital libraries are] human activity systems that unite readers, authors, librarians and researchers with electronic materials, resource streams, computer equipment and know-how (Covi & Kling, 1996).

We think it best to consider a *digital library* to be a library / information service, located either in a physical or virtual space, or a combination of both, in which a significant proportion of the resources available to users of that service exist only in digital form. This is a pragmatic decision, taken so as to include within its scope all significant research under the heading of digital libraries. It also covers approaches described by other terms, which will now be considered.

The *electronic library* is a term which has had a longer usage in the literature than 'digital library', and is usually associated with a somewhat old-fashioned approach. It generally indicates a rather limited approach to the digital library, simply indicating the provision of a range of material in digitised form, within the framework of traditional library provision.

The *hybrid library* is generally taken as lying somewhere on a continuum from the traditional to the digital library, with electronic and paper-based sources used alongside one another. Sutton (1996) who saw it as lying on the continuum traditional-automated-hybrid-digital originally introduced the term; Oppenheim and Smithson (1999) describe the history of the term, and its usage.

The challenge of the hybrid library is to integrate access to sources in a variety of formats, and from both local and remote sources. Its proponents regard it as a worthwhile model in its own right, rather than simply a transition state between traditional and digital libraries (Pinfield et al. 1998, Brophy & Fisher 1998, Oppenheim & Smithson 1999). *The gateway library* is a term with much the same meaning (Dowler 1997).

A hybrid, or gateway, library provides an environment and services that are partly physical and partly virtual. Though it assumes an ideal of closer integration of information provision, regardless of format and medium, it does not assume, as the full digital library does, that the migration of all materials to digital form is either feasible or desirable:

> After a decade of messianic expressions like 'Electronic Library' and 'Digital Library', which seemingly imply there is no longer a place for books, the term 'Hybrid Library' can be seen as evidence of a refreshing, new realism' (Price 1998)

Crawford (1999) refers to this type of library as a *complex library*. This seems to us to be a particularly apposite term, as we discuss in a following paper (Rowlands & Bawden 1999).

The term *virtual library* has been used at times, though with little consistency in meaning; it is often used to describe collections of web resources. *Library of the future* seems to be a catch-all term, while *library without walls* has sometimes been used to refer, not only to digital collections, but also to outreach programmes with physical material. None of these terms seems particularly useful.

As we have seen, the term *digital library* has been used, somewhat confusingly, to subsume all the concepts noted above. It does, however, usually carry two particular connotations: a greater *degree* of digitisation, i.e. a greater reliance on digital sources, and a greater ambition and scope.

The way in which the concepts described by these various terms can be encapsulated in conceptual models of the digital library is discussed in a later paper (Rowlands & Bawden 1999).

Associated concepts

Documents and collections

It is clear that one attribute of any digital library is a diversification to encompass different material types and formats, raising issues of integration as well as cataloguing and navigation (Davies 1998). Nonetheless, it is still possible to think in terms of the concepts of documents and collections; though these may have different meanings in the digital environment.

The *documents*, or resources, in a digital collection will be of very varied nature, as indeed are physical documents. They can be understood as single digital objects, with the following characteristics

- unique: they should be identifiable as the same resource in whatever format or medium they are instantiated;
- coherent: they must present a logically coherence quantity of information;
- significant: they must include a viable and useful information content;
- access: they must be accessible by the libraries' systems.

Commentators have suggested various topologies of resources. Rusbridge, for example, proposes that they can be usefully categorised as legacy, transition, new and future – but there is no agreement as to the most sensible such categorisation. From a pragmatic point of view, it seems that that a simple three-way categorisation is useful:

- non-digital: resources not in a digital form, and which will not be converted in a foreseeable future, but which will be managed and controlled by digital means;
- transitional: resources originally produced in non-digital form, which have been subsequently digitised;
- digital: resources designed and created in digital format.

(The term 'non-digital' is somewhat negative, but is preferable to 'legacy', since the latter is has a strong negative connotation of obsolescence.)

The definition of *collection* in a digital environment has to be adapted to allow for the distributed and dynamic nature of the digital library (Jones 1999). Collection membership is based on 'criteria rather than containment': resources are members of a collection by virtue of conforming to a set of formal criteria, e.g. subject classification (Logoze & Fielding 1998). This implies that a collection may be created and varied in a dynamic, perhaps even in an automatic, way based either on the content within the resources or on metadata describing them. The lifecycle of collections may then be entirely independent of the objects composing them; some collections will be effectively permanent, while others may be created according to short-term needs.

A digital library collection, which, as noted above, may be logically defined as a set of criteria for selecting resources from a broader information space, may be less formally understood as a set of digital objects, with the following characteristics

- coherence: there is some theme to the set, it is not a random grouping;
- significance: it is useful enough in sum to be worth regarding as a collection;
- control: the objects are prepared and organised according to some standards, and are preserved in some way;
- access: there must be means of access to the objects, usually both by description and subject.

Other criteria can be suggested; for example, Crawford (1999) includes uniqueness, since a set of commonly available objects will not generally be regarded as a collection, but the four above will generally be sufficient to define a digital collection.

This leaves, however, a difficulty in placing a clear boundary around a collection. If, for example, an item selected for a collection is an HTML (HyperText Markup Language) file, it may have links to other HTML documents that the user can freely access. Are these also a part of the collection? What about files which link *to* the file in the collection? These, and similar questions, which have no clear answers, may have important legal implications (Logoze & Fielding 1998).

In a traditional library, the collection is defined by physical presence and by ownership. In the digital library, its must be defined by access – technical and economic – since the items comprising the collection will not be physically present, and will mostly be licensed rather owned outright. The impact of this on the nature and economics of the acquisitions system should not be under-estimated:

> ... the terms of licenses in the electronic world are limited only by the imaginations of corporate lawyers. There will be a variety of terms and conditions, a variety of pricing models, and all possibilities are likely to exploited. Given the likely huge number of licensed digital objects (eventually millions) and the large numbers of copyright owners (many, many thousands) this could result in an absolutely untenable situation for libraries struggling to achieve compliance and to provide a useable service to their patrons (Rusbridge 1998).

Digital, electronic, multimedia

These terms – digital, electronic, multimedia – are used with abandon, and little precision, in the literature. It is most appropriate to think simply of digital materials, noting that their interconversion and media requirements, while in

principle straightforward, may in practice be problematic. Note also that there is a significant difference between digital objects created in electronic form, and hence amenable to digital processing, and those that have been digitised from printed form. The latter may or may not be so amenable; e.g. a printed book may be digitised by keyboarding, or by scanning plus OCR (optical character recognition), so that its text is searchable, or it may simply be digitised by scanning so that it can only be viewed.

Conclusions

There is little agreement on the basic assumptions underlying the concept of the digital library, to judge from the literature of the subject. Nor is there general agreement about the meaning of some of the most essential terminology, even 'digital' itself. This is not merely a problem of academic hair-splitting; it affects the ability of funding agencies, researchers and practitioners to reach a common understanding of aims and objectives, and hence compromises the successful future development of library services.

One way of dealing with these difficulties is to attempt to establish a wide-ranging conceptual framework, able to assimilate the variant understandings of terminology, and to allow a consensual understanding to develop. This is a logical next step in this research study. The attempt at establishing such a framework is addressed in another paper by the authors (Rowlands and Bawden 1999).

References

Aparac, T. et al. (eds.) 1999. Digital Libraries: interdisciplinary concepts, challenges and opportunities, Proceedings of the 3rd CoLIS Conference, Zagreb.

Bawden, D. 1999. Information literacy, digital literacy, and other literacies; a review of concepts and models, submitted for publication.

Bawden, D. and I. Rowlands. 1999. Understanding digital libraries: towards a conceptual framework. British Library R&I Report No 170.

Borgman, C. L. 1999. What are digital libraries? Competing visions. *Information Processing and Management* 35(3): 227–43.

Brophy, P. 1999. Digital Library Research Review. Library and Information Commission Report No. 17. London.

Brophy, P. and S. Fisher. 1998. The hybrid library. *New Review of Information and Library Research* 4:3–15.

Buttenfield, B. 1999. Usability evaluation of digital libraries, in D Stern (ed.) Digital libraries: philosophies, technical design considerations, and example scenarios. Binghampton, NY: Haworth Press, 39–60.

Chowdury, G. G. and S. Chowdury. 1999. Digital library research: major issues and trends, *Journal of Documentation*, 55(4): 409–48.

Covi, L. and R. Kling. 1996. Organizational dimensions of effective digital library use: closed rational and open natural systems models. *Journal of the American Society for Information Science 47(9)*: 672–89.

Crawford, W. 1999. Being analog: creating tomorrow's libraries. Chicago, IL: ALA Editions.

Cronin, B. 1998. Information professionals in the digital age. *International Information and Library Review* 30(1): 37–50

Davies, C. 1998. Future user issues for the networked multimedia electronic library. ELINOR: Electronic Library Project. New York, NY: Bowker-Saur, 105–30.

Dowler, L. 1997. Gateways to knowledge, the role of academic libraries in teaching, learning and research. London: MIT Press.

Elliott, M. and R. Kling. 1997. Organisational useability of digital libraries. *Journal of the American Society for Information Science* 48(11): 1023–35.

Ercegovac, Z. 1997. The interpretations of library use in the age of digital libraries. *Library and Information Science Research* 19(1): 35–51.

Esler, S. L. and M. L. Nelson. 1998. Evolution of scientific and technical information distribution. *Journal of the American Society for Information Science* 49(1): 82–91.

Garrod, P. 1999. Survival strategies in the Learning Age – hybrid staff and hybrid libraries, *Aslib Proceedings* 51(6): 187–94.

Gilster, P. 1997. Digital literacy. New York: Wiley.

Hayes, M. 1997. Get it right first time: a beginner's guide to document management. *Information Management & Technology* 30(2): 84–85.

Hey, J. M. and A. Wissenburg. 1998. Modelling the hybrid library. *New Review of Information and Library Research* 4: 103–10.

Jones, D. 1999. Collection development in the digital library, in D. Stern (ed.) Digital libraries: philosophies, technical design considerations, and example scenarios. Binghampton, NY: Haworth Press, 27–38.

Kuny, T. and G. Cleveland. 1998. The digital library: myths and challenges. *IFLA Journal* 24(2): 107–13.

Lang, B. 1998. Developing the digital library, in Towards the Digital Library: the British Library's Initiatives for Access programme. London: British Library, 227–33.

Lesk, M. 1997. Practical digital libraries. San Francisco: Morgan Kaufman.

Levy, D. M. and C. C. Marshall. 1995. Going digital: a look at assumptions underlying digital libraries. *Communications of the ACM* 38(4): 77–84.

Licklider, J. C. R. 1965. Libraries of the future. Cambridge, Mass.: MIT Press.

Logoze, C. and D. Fielding. 1998. Defining collections in distributed digital libraries. *D-Lib Magazine* (November).

Odlyzko, A. 1997. Silicon dreams and silicon bricks; the continuing evolution of libraries. *Library Trends* 46(1): 152–67.

Oppenheim, C. and D. Smithson. 1999. What is the hybrid library? *Journal of Information Science* 25(2): 97–112.

Pedley, P. 1999. Intranets and push technology: creating an information-sharing environment. London: Aslib.

Pinfield, S. et al. 1988. Realising the hybrid library, *New Review of Information Networking* 4: 3–21 [also at *D-Lib Magazine*, October 1998, URL [http://mirrored.ukon.acuk/lis-journals/dlib/dlib/dlib/october98/10pinfield.html]

Price, D. J. 1998. The hybrid library and collection development. *New Review of Library and Information Research* 4: 129–39.

Rowlands, I. and D. Bawden. 1999. Digital Libraries: a conceptual framework. *Libri* 49(4): 192–202.

Rusbridge, C. 1998. Towards the hybrid library. *D-Lib Magazine* (July/August). URL: http://www.dlib.org/dlib/july98/rusbridge/07rusbridge.html

Saracevic, T. 1999. Preface. Digital Libraries: interdisciplinary concepts, challenges and opportunities, Proceedings of the 3rd CoLIS Conference, edited by T. Aparac et al. Zagreb, xii.

Schatz, B. R. 1997. Information retrieval in digital libraries: bringing search to the net. *Science* 275(5298*)*: 327–34.

Smith, M. 1998. The next five years: World Wide Web, Internet, and .? In Interlending and Document Supply: Resource Sharing Possibilities and Barriers, Aarhus, Denmark, 24–28 August 1997. The Hague: IFLA: 179–96.

Stern, D. 1999. Digital libraries: philosophies, technical design considerations, and example scenarios. Binghampton, NY: Haworth Press.

Sutton, S. 1996. Future service models and the convergence of functions. In K. Low (ed.), The roles of reference librarians, today and tomorrow. New York: Haworth Press, 125–43.

Travica, B. 1997. Organisational aspects of the virtual/digital library. *Proceedings of ASIS* 34: 149–61.

Appendix: Concept analysis

Use of 'digital library' terms over time

In order to assess the relative occurrence of various terms related to digital libraries, the *Social Sciences Citation Index*® and *LISA* databases were searched for the terms 'digital library / libraries', 'electronic library / libraries', and 'hybrid library / libraries', in the titles, abstracts or assigned index terms of records. The results are ranked by year of publication. While somewhat rough and ready, this approach gives a general indication of the variation of usage over time

	Digital library		Electronic library		Hybrid library	
year	LISA	SSCI	LISA	SSCI	LISA	SSCI
1998	152	48	78	18	4	–
1997	197	57	153	40	3	–
1996	129	37	178	27	–	–
1995	67	27	188	36	–	–
1994	12	2	106	24	–	–
1993	9	4	57	12	–	–
1992	3	2	31	10	–	–
1991	2	–	28	7	–	–
1990	1	2	20	7	–	–
1989	–	–	10	6	–	–
1988	–	–	12	7	–	–
1987	–	–	11	6	1	–
1986	–	–	9	5	–	–
1985	–	–	7	9	–	–
1984	–	–	5	16	–	–
1983	–	–	10	3	–	–
1982	–	–	8	1	–	–
1981	–	–	12	7	–	–
1980	–	–	–	–	–	–

These results indicate that: 'hybrid libraries' have little literature representation; 'electronic libraries' first appeared in the literature pre-1980, and have been consistently represented since, with peaks in the early-mid 1980s and the mid-late 1990s; and 'digital libraries' first appeared in the literature in 1990 and occurrence has grown rapidly through the decade.

AltaVista term clustering analysis

A great deal of information on digital libraries is accessible from the Internet. This is very varied in nature, provenance and quality, as with much Web-based information. A simple analysis of this corpus can shed some light on the way in which the various terms are used, and the concepts associated with them. This was achieved for the purposes of this study by searching the AltaVista search engine for several relevant phrases, and submitting the results to the *refine* process. This carries out a statistical analysis of word co-occurrence, clusters the results into sets of related items, and denotes these by the 'typifying' words for each cluster. Although this is a rather unsophisticated approach, it serves to give a rapid overview of the way in which the concept searched for is 'viewed' on the Internet.

The results are shown here for three search phrases: digital library; electronic library; hybrid library. Varying the form of search phrases gave no additional useful information. The percentage figures indicate the proportion of the output assigned to each cluster; since some items are not assigned to any cluster, while others may be assigned to more than one, the sum of the percentages will not be 100%.

Digital library

54% libraries, digital, library, librarians, collection, collections, historical, Congress
40% metadata, Dublin, attribute
32% schoolkids, kids, resources, educational, sites, resources, links, yahooligans
28% testbed, DARPA, infrastructure, distributed, project, architecture, object
27% retrieval, indexing, semantic, database, ACM, IEEE
26% scholarly, ariadne, electronic, journals, interlibrary, publishing, humanities
21% searchable, site, annotated
19% cataloguing, librarian, OCLC
19% cnri, repository, repositories
14% engines, searches, browsing, searching, engine, search
12% teachers, educators, science, classroom

Electronic library

76% library, libraries, librarians, electronic
31% selector, merit, funded, administered, sponsored, Michigan
28% searchable, indexed, lists
23% cataloguing, retrieval, bibliographic, collection, materials, OCLC
21% databases, journals, newspapers, journal, articles, titles, database
20% periodicals, reference, periodical, abstracts, citations, indexes, publications
16% sites, links, resources, information, virtual, government, federal, school
12% gopher, gophers, telnet
9% humanities, science, sciences, social, university, research

Hybrid library

40% hybrid, CD-ROM, hybrids, library, algorithms, electronic, database
16% digital, digitization, preservation
9% bandwidth, ATM, switching
7% genes, genome, YAC

The clustering for ‘digital library’ shows the greatest richness of concepts. The general digital library cluster (1), is complemented by concepts of retrieval (5,7,10), publishing and subject access (6,8), metadata (2), repositories (9), educational applications (3,11), and research projects (4).

‘Electronic library’ shows a somewhat more ‘old fashioned’ picture, with concepts of libraries (1), and rather traditional looking processes (3,4) and resources (5,6), as well as rather dated Internet tools (8).

‘Hybrid library’ shows an overlap with the semantically unrelated idea of libraries of hybrid organisms in genetics; apart from this, it shows a concern for the processes of digitization and the subsequent preservation of material (2) above general ‘electronic library’ issues (1).