

Law, Derek (2010) Digital Libraries in Higher Education. In: Business Planning for Digital Libraries: International Approaches. Leuven University Press, Leuven, pp. 33-43. ISBN 9789058678379

<http://strathprints.strath.ac.uk/26637/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<http://strathprints.strath.ac.uk>) and the content of this paper for research or study, educational, or not-for-profit purposes without prior permission or charge. You may freely distribute the url (<http://strathprints.strath.ac.uk>) of the Strathprints website.

Any correspondence concerning this service should be sent to The Strathprints Administrator: eprints@cis.strath.ac.uk



3

DIGITAL LIBRARIES IN HIGHER EDUCATION

Derek Law

Vision and mission of the digital library

Higher Education libraries tend rather thoughtlessly to be considered a necessary, if expensive, part of a university which requires little justification. By extension the same approach often characterizes digital libraries. And yet without a clear understanding of the purpose of such libraries, expensive white elephants can all too easily be built. Visions and missions for digital libraries are quite rare. Partly because, in the absence of any common view of the nature of the digital library in Higher Education, it is then important to clarify why a digital library is being created. A surprisingly large number of digital libraries appear to have neither vision nor mission, at least according to their websites. This may be due to the fact that they are embedded as part of their larger institutional libraries. Imperial College, Cambridge University and Columbia University, for example, are all silent on their large digital library programmes. Others are more forthcoming and usefully illustrate the differences in what is being attempted. Perhaps the most ambitious, but certainly the clearest, mission statement comes from the British Library (2009): “The Digital Library Programme’s mission is to enable the United Kingdom to preserve and use its digital output forever.”

The vision is equally crisp:

“Our vision is to create a management system for digital objects that will

- ingest, store and preserve any type of digital material in perpetuity
- provide access to this material to users with appropriate permissions
- ensure that the material is easy to find
- ensure the authenticity of the material
- ensure that users can view the material with contemporary applications
- ensure that users can, where possible, experience material with the original look-and-feel”

Other definitions are either fairly generic or concerned with process. The Oxford Digital Library (ODL) has Principles and Guidelines which commit it to “build up a significant set of digital resources for local and remote online access. Like traditional collection development, long-term sustainability and permanent availability are major goals for the ODL. Therefore the use of standards in the digital conversion process and for the description of digital resources will be essential for projects funded through and supported by ODL” (Oxford, 2009). Harvard notes that its “Initiative was consciously constructed on a different model: the integration of digital resources into the existing



Derek Law

library structure. Integrated access to the collections, regardless of format, was a key aim of the Library Digital Initiative” (Harvard, 2009). At Hong Kong University Library “more digital projects are being developed to provide continuous access to digital content and services” (Hong Kong, 2009). And finally the Glasgow Digital Library provides “a distributed digital library based in Glasgow which aims to produce a coherent digital learning and information environment for Glasgow’s citizens, through the development and implementation of a common collection development policy and an agreed technical and inter-working infrastructure” (Glasgow, 2009). One of the longest standing digital libraries is the California Digital Library, explored more fully elsewhere in this volume where its origins and mission are fully described. As far as its current public presence on its website goes, however, it “supports the assembly and creative use of the world’s scholarship and knowledge for the University of California libraries and the communities they serve” (California, 2009). In addition, the CDL “provides tools that support the construction of online information services for research, teaching, and learning, including services that enable the UC libraries to effectively share their materials and provide greater access to digital content.”

This lengthy set of examples is intended to demonstrate that there remains a degree of vagueness in public statements as to why digital libraries are being created. This is understandable in some cases where these libraries were early explorers and adopters of the digital world. However, after the project and exploration stage there is a real need for clarity of purpose if the case for creating a digital library is to be made and accepted. Even rarer than the sort of vision and mission described above is a link to the institutional strategic plan. And yet if the digital library is to have a context and if it is to have institutional support, it should be obvious that it must be part of delivering the overall aims and objectives of the institution.

The business case

Interestingly, none of the visions described above considers the business case for investing in digital libraries. University libraries do not stand in isolation, although all too often they appear isolated – or at least insulated – from the rest of the institution. If the business of the university may be loosely defined as teaching and research, the business case for creating digital libraries should relate to how the business of the university is supported, promoted and ideally enhanced. Tenopir (2009) has begun some work on this area, and while the evidence thus far is qualitative rather than quantitative it does have regard for how digital libraries promote research. Allowing researchers to cut out lines of research already used by others, exploring what will assist grant applications, identifying other research teams for joint research proposals, exploring interdisciplinary boundaries in novel ways, gaining access to raw research data and so on form the basis for a powerful business case built around the general well-being of the institution rather than the exclusive well-being of the library. Even the well documented increase in the number of articles read by researchers may form part of a justification (Tenopir et al, 2009). A recent CIBER (2009) study also shows massive increases in the use of electronic resources. Journals are the lifeblood of research. There is a strong correlation between e-journal use and research outputs. The ROI (Return On Investment) can reasonably be demonstrated to lie in how researchers perform rather than simply in how the library is used.



Higher education libraries have not hitherto had to undertake a great deal in the way of financial planning. In the United Kingdom, for example, the budget is typically last year's figure plus a few per cent. Until very recently the Library was seen as a necessary, if expensive, part of the fabric of any university. Of course the budgets were and are very well managed, but little was needed in the way of business planning and such revenue generation as was undertaken either tended to be for endowments or was a way of paying for new services the costs of which were readily identifiable, whether photocopying, interlending or online searching. In essence the Library was simply a top-sliced cost from the University budget, or delegated to faculties, as is often the case in continental European universities.

There are more complicated methods of calculating value, using economic models. Attempts have been made to put a monetary value on library activity (Baldwin, 2004) as a way of demonstrating value for money, while the British Library (2004) used contingent valuation to demonstrate that it creates value at a rate of 4.4 times the level of its grant. Kelly (2008) has estimated that the value of knowledge transfer from university libraries in Scotland to the external community can be calculated as £7.1 million. This calculation is part of a wider study on the economic impact of universities, which identifies the market value of activities, in this case using nationally collected statistics on the number of visits made to university libraries in Scotland by external users.

However there is a wider issue than identifying costs and attributing them or charging for services. It may well be unhelpful to try and isolate the value and costs of the digital library. University libraries are part of the wider academic exercise, and not simply a cost. Thus the existence of a great library may be a factor in student recruitment, for example, but this would not necessarily show up in a simple calculation. Roosendaal (2003) makes a key point:

The economic impact of ICT on the academic institution and its library cannot be discussed in an isolated way...the costs of establishing a digital academic library should be considered as part of the integral costs of establishing an overall digital environment for the entire academic institution.

Making the business case for a digital library is then much more than simply identifying the costs of a new service and attributing income.

Target user group(s)

It is usually straightforward to define the user groups in higher education as either members of the university or some subset of them such as students. In many cases a digital library begins by undertaking learning tasks such as digitising examination papers and providing a repository for these. These are clearly aimed at the student community, as is more general support for a digital learning environment. The purchase of access to e-journals will be targeted more clearly at researchers.

Most digital libraries focus on content, but the fruitful area of services also needs to be considered. Training can be aimed at all groups, while identifying new research tools such



Derek Law

as Openwetware (2009) or Blue Obelisk (2009) is much more clearly targeted at the research community. Twenty-four hour reference services can also be facilitated by the use of networks (Davis, 2004) and, even if little used, can be symbolically important.

User design principles

The key but often neglected element of digital libraries is that they must aim to be integrated into the workflow of users. For example, availability is important. We know from weblogs that 25% of use is outside the traditional working day and 15% is at the weekends (Nicholas, 2009). There is also a need for simplicity. Advanced search tools are used by a vanishingly small number of users and are simply not worth the effort of inclusion (Nicholas, 2009). It would also be valuable to consider how far the library might help to meet new requirements placed on the university in a cost-effective way. Thus research funders, for example, increasingly require the research team to maintain project websites for some years after funding ends. They may also require a data curation plan; they mandate depositing research outputs in repositories, and a growing number of institutions now require this too. It should be possible to demonstrate that all of these activities fall under the broad heading of the organisation of knowledge, which is the cornerstone of what information science and libraries are about. We also know that users wish to be able to speed through the site, and (perhaps counter-intuitively) the best researchers speed through the fastest (Nicholas, 2009). Speed rather than comprehensiveness should be a key design aim. Coupled with this is familiarity. There is resistance to learning another system when the system you have invested learning time in works well enough. Where possible the digital library should mimic existing popular services. And, if possible, another aim should be the ability to customise services for individual users rather than building the one size fits all platforms which we have tended to do hitherto.

Collaboration and shared infrastructure can assist in reducing costs and adding value, and should be explored. NARCIS (2009) provides access to all Dutch science, almost 250,000 scientific papers as well as datasets, with the vast bulk available on open access. The Dutch university system with only 13 universities is perhaps more easily organised than many, but regional or disciplinary consortia can also follow this approach.

Technical approach: architecture, infrastructure, metadata

Although every case is different, at the extremes there are two kinds of digital library. The first and perhaps commoner is a reflection of the past and perhaps arises from a different era where the Library was a taken for granted common good. For a decade most higher education librarians have preferred to ignore the mushrooming growth of born digital material and focus on the past. They have created consortia to negotiate deals with publishers, but deals which are usually declared unsatisfactory, managing the commercial output of the journals industry. They have learnt the uncomfortable truth that increasing access to digital material comes at the price of leasing it rather than owning it in many

cases. At the same time we have generally preferred to digitise the paper collections we already own, rather than dealing with the born digital material our staff and students are creating. At the other extreme lies what may broadly be called the repository movement, which does aim to address institutional outputs. The University of Adelaide, for example, with its *Adelaide Research & Scholarship* project offers a sort of hybrid which “provides a platform for the collection, organisation, access, annotation and preservation of information in digital formats, as well as digital management of information in physical formats. Its primary focus is on the scholarly output of members of this University, and items of interest to those members (for example the rich resources of Special Collections)” (Adelaide, 2009).

It is not possible for every library to support and manage all of its ‘business processes’, especially as the demands on the library grow, and service expectations and technologies change so quickly. Libraries have historically depended on shared platforms for services, and we may be about to see another step change in adoption. The motivation is to remove redundancy and to build capacity through collaboratively sourcing solutions, so as better to focus library effort on where it can make a distinctive local impact on the quality of the research and learning environment (Dempsey, 2006). The ideal then is perhaps to aim for shared common infrastructure, but to take sole responsibility for the bibliographic integrity and metadata of what is produced locally. This may imply a distributed architecture as we may expect researchers, for example, to wish to hold their own datasets rather than pass them on to the library. This may not matter if the library is seen as the arbiter of policy and standards.

Costing

It is a curiosity of the traditional library that we know very little about its real costs. We know a great deal about budgets, and thanks to long-standing arrangements for data collection we have excellent time series analysis of how patterns have changed. But we know almost nothing about the life-cycle costs of libraries and their indirect costs. And yet it is very easy to demonstrate that the costs for buildings and utilities alone are probably greater than the typical annual library budget. So there are significant dangers in attempting to compare the costs of physical libraries with those of digital libraries. It is still rare for a digital library not to be part of a physical library.

How does one assign costs and value? Perhaps the biggest mistake is to do it in the context of the library alone. It is possible to do this mechanistically as the major LIFE (Life Cycle Information for E-Literature) Project demonstrated. The project is based at University College London (McLeod, 2005) and it aims at a complete analysis of all the activities which relate to the management of digital content, from its selection, through licensing and acquisition to ingest, metadata creation, adding links, access, user support, storage costs and preservation. The study used evidence from large scale examples to calculate average costs. This study then provides at least the basis for creating a very robust methodology to define digital library costs. The project has made some quite specific estimates of costs, based on British experience:



Derek Law

It established that in the first year of a digital asset's existence:

- The lifecycle cost for a hand-held e-monograph is £19
- The lifecycle cost for a hand-held e-serial is £19
- The lifecycle cost for a non hand-held e-monograph is £15
- The lifecycle cost for a non hand-held e-serial is £22
- The lifecycle cost for a new website is £21
- The lifecycle cost for an e-journal is £206

LIFE also predicts that in the tenth year of the same digital assets' existence:

- The total lifecycle cost for a hand-held e-monograph is £48
- The total lifecycle cost for a hand-held e-serial is £14 per issue
- The total lifecycle cost for a non hand-held e-monograph is £30
- The total lifecycle cost for a non hand-held e-serial is £8 per issue
- The total lifecycle cost for a new website is £6,800
- The total lifecycle cost for an e-journal is £3,000

It is interesting to note that the precision of these numbers is almost impossible to match for our print on paper libraries. We should also understand that most of these costs will apply to non-commercial material. In all the debate about economics the huge volumes of grey literature, donations and archives which research libraries acquire each year tend to be neglected. And these costs are just as real for their electronic equivalent in digital libraries.



The common view appears to be that digital libraries cause increased costs (Baker & Evans, 2009). The authors list the cost of digitising material, the cost of copyright clearance, constant technological change and innovation, training, preservation, administration and so on. But it is possible to consider not only the additional costs which a digital library will incur but the savings which may be made elsewhere. Ayris (2005) has suggested that savings in inter-lending costs could be used to fund a repository. Law (2009) has noted that at Strathclyde the cost of utilities to support a million volume library was £500,000 in 2007/8. By investing in electronic materials and disposing of paper stock Strathclyde plans to halve the physical space the library will occupy. Crudely this will save half the utility bill, but a more substantial saving will come if the university is able to shed a teaching building in the city centre and use the saved library space for teaching rooms. This reduction in the university estate will have even greater savings. We know a huge amount about the direct costs of traditional libraries, but almost nothing about their indirect costs and overheads. These could well look very different for digital libraries, not least because some of the costs, such as equipment, move downstream to the user and away from the university.



Another interesting, if largely unexplored, area is that of shared services and open source systems. Library automation systems are expensive to purchase and have high recurrent costs. As we move on, there is renewed interest in creating shared services. These are now seen as utility-like functions where costs can be reduced by sharing. There is also interest in open source software, whether for repositories – such as D-Space, Greenstone and ePrints – or for library management systems such as Koha, OpenBiblio and PhPMyLibrary. One of the curiosities of Open Source is that university procurement offices can have difficulty accommodating this, since there is no one to tender in a conventional way.



Anticipated income streams

Lesk (2004) has perhaps done most work on trying to identify how digital libraries might be funded. A traditional British University Library consumes about 3% of institutional income and produces very little income beyond fines and photocopying.

Of that sum rather more than half will go on staffing, one third will go on acquiring content and the remainder will be spent on computing services, furniture and supplies.

Higher education libraries tend to indulge themselves in a myth that they are free. This has never been true. Although they are generally funded through some form of top-slice of university budgets, a range of services has always been charged for and/or other services are rationed against a notional charge. Thus services as varied as binding and inter-lending may be charged for. Photocopying is usually charged; on-line searching was often a charge. And some charges such as fines are in practice often in reality used as a *de facto* lending fee by at least some students. And, of course, external users (such as health service employees) are often charged for individually or in bulk.

Nevertheless, Collier's (2004) work on TEL (The European Library) resulted in the business case resting on providing value for money rather than on considering how income can be generated, and based this on several criteria:

- That digitised content would be important
- That the library should concentrate on material not readily available elsewhere
- That material should be as far as possible free at the point of use
- The strong public service ethic of the participating national libraries

Interestingly this work focused on content and not on services, which are a much neglected area of digital libraries, but most higher education libraries would feel comfortable with this approach. Subsidy from the parent institution will almost certainly prove to be the default mode of funding. A now somewhat dated but still valid study of Digital Library Federation members in 2001 (Greenstein and Thorin, 2002) showed that "ordinary" library funds were being diverted into support for digital library activities. A range of other models has been proposed:

- Subscription: password-based with the password having a fee
- Pay per use: micropayment systems would pay for this and we know that there is available money in the system from experience with photocopying and printing.
- Licensing: it is less clear whether this would be charged directly or as a kind of lab fee.
- Author/owner pays: experiments with this are inconclusive so far
- Sponsorship: there is little evidence that this would prove viable, not least at a time of great economic stringency.
- Advertising: this is an environment to be explored. However, early experiments have proved discouraging in terms both of the level of income generated and the ambition of various bodies to have powers of veto over what may be advertised.

Derek Law

Marketing

There is a certain irony in the fact that the more transparent and easy to use a digital library becomes and the more integrated it becomes with network tools such as Google or larger resources such as Amazon, the less visible and obvious it becomes to users. This is particularly important with senior academic staff, as increasingly those who make decisions about the size and shape of library budgets are the least likely to visit and use the library as a place. Thus marketing the library's services as well as its content becomes a critical activity. Publishers have brands which they work hard to maintain. Very few libraries, other than great historic ones such as the Bodleian Library, have brands. Marketing therefore becomes critical. The one thing we have learned above all from digital services, products and environments is that the market will decide, not quality or price or even availability. It is vital to be responsive to user needs and to recognise that the user has alternatives.

Risk analysis

Some of the risks to the digital library are obvious, and to an extent generic. Power failure, flood, fire and theft all have huge potential to damage. But perhaps the area which requires most management is the whole area of access to content. The loss of access to leased data through financial problems, the loss of access to data through government prohibition or failure of companies, the balance between locally hosted and externally accessed data, and LOCKSS (2009) type arrangements to ensure data security all require careful consideration. Other risks may be associated with staffing problems, the failure of organizational partner agreements, key stakeholders not buying into or supporting the project, technical hardware and software issues. As always, the best strategy is to spread the risk. No one member of staff should be critical, there should be multiple partners, there should be alternative or overlapping stakeholders and the use of generic hardware and software will allow easy replacement. Most basically of all, creating, maintaining and monitoring a risk register is now considered basic good practice.

Implementation plan

Implementation plans are *de rigueur* for all major projects. Again the key is not so much the creation of the plan but its monitoring and maintenance. There is also a danger that such plans can be monolithic rather than responsive. Ideally the plan should recognise the importance of a feedback loop, so that system design responds to users. One substantial tool which is available is log analysis. Surprisingly little attention is paid to this rich source of information.



Financial planning

A financial plan is much more important for a digital library than a conventional one. If nothing else, the experience of working with library automation systems shows the need for a proper understanding of equipment replacement cycles. Brewer (2002) describes a service at the University of Derby which did not prove sustainable, although it was very popular. This emphasises the need for the business plan to reflect the needs of the institution, and in turn for the financial plan to be realistic about where the emphasis of investment should be to provide a service tailored to the needs of the institution. Concepts such as break-even point and ROI may not be entirely quantifiable, as discussed above, but there should certainly be a clear understanding of when and how the digital library will reach some kind of equilibrium. There will undoubtedly be new and additional costs, and these should be clearly understood and budgeted for. If other costs (such as interlending) are to be netted off against these new costs there needs to be a shared clarity of what this means.

Conclusions

The economics of digital libraries in higher education remain at a very primitive stage, not least because our lack of understanding of the overall costs of traditional libraries is so incomplete. Many digital libraries appear to emerge as services from projects without a clear understanding of their role, function and cost. Yet these are the most important factors. Clear articulation of a vision and mission are essential pre-requisites for the creation of the business case, technical design and service definition which must lie at the heart of the development.

Summary

- Visions, missions and business plans for DLs are quite rare in Higher Education and the understanding of the economics is at a primitive stage.
- Business planning and business cases may be essential in a way in which they were not for traditional libraries
- The business case should ideally be based on the institutional strategy
- User design and responsiveness are even more important for DLs than for traditional libraries as students and researchers change their habits
- Planning should encompass two distinct aspects: access to the published sources and access to/preservation of institutional research or digitised content
- Costing should cover total costs of ownership. It is not safe to assume that the total cost of DLs is higher than that of traditional libraries over time.
- Focus on services may open up new income streams
- As the institutional DL gets covered by global web services more local marketing and PR may be needed to demonstrate value and gain credit for the initiative





Derek Law

References

- Adelaide (2009) <http://digital.library.adelaide.edu.au/> (viewed 5 June 2010)
- Ayris, Paul (2005). Note of Research Communications Forum One-day Conference – 7 March 2005 www.berr.gov.uk/files/file10867.doc (viewed 5 June 2010)
- Baker, David & Evans, Wendy (2009) Digital Library Economics: the key themes. In Baker, D. & Evans W. Digital Library Economics: an academic perspective. Cambridge, Chandos.
- Baldwin, Jerry (2004) Mn/DOT Library Accomplishments. *TRUpdate* 9, Spring
- Blue Obelisk (2009) http://blueobelisk.sourceforge.net/wiki/Main_Page (viewed 5 June 2010)
- Brewer, Gordon, (2002) The University of Derby Electronic Library: a case study of some economic and academic aspects of a local digitised collection. Program: electronic library and information systems vol 36 pp.30 - 37
- British Library (2004) Measuring Our Value. <http://www.bl.uk/pdf/measuring.pdf> (viewed 5 June 2010)
- British Library (2009) <http://www.bl.uk/aboutus/stratpolprog/digi/dom/mission/index.html> (viewed 5 June 2010)
- California Digital Library (2009) <http://www.cdlib.org/> (viewed 5 June 2010)
- CIBER (2009) E-journals: Their Use, Value and Impact London:RIN
- Collier, M (2004) Development of a business plan for an international co-operative digital library - The European Library (TEL). *Program* 38, no. 4 pp225-231
- Davis, K. & Scholfield, S. (2004) "Beyond the virtual shore": an Australian digital reference initiative with a global dimension *Library Review* 53, pp61-65
- Dempsey, L. (2006) The (Digital) Library Environment: Ten Years After. *Ariadne*, No 46 <http://www.ariadne.ac.uk/issue46/dempsey/> (viewed 5 June 2010)
- Glasgow (2009) <http://gdl.cdli.strath.ac.uk/documents/gdloverview.htm> (viewed 5 June 2010)
- Greenstein, D. & Thorin, S. (2002) The Digital library: a Biography. Digital Library Federation, Washington, DC
- Harvard (2009) <http://hul.harvard.edu/ois/ldi/> (viewed 5 June 2010)
- Hong Kong (2009) <http://lib.hku.hk/database/> (viewed 5 June 2010)
- Kelly, Ursula, McNicoll, Iain & Brooks, Richard. (2008) Towards the estimation of the economic value of the outputs of Scottish HEIs: Next Steps Project. Final Report. Glasgow, University of Strathclyde <http://eprints.cdli.strath.ac.uk/3106/> (viewed 5 June 2010)
- Law, D. (2009) An awfully big adventure: Strathclyde's digital library plan <http://www.ariadne.ac.uk/issue58/law/> (viewed 5 June 2010)
- Law, D. (2009a) Digital Library Economics: Aspects and Prospects in Baker, D. & Evans, W. Digital Library Economics. Oxford, Chandos, 2009
- Lesk, Michael (2004) in Andrews, J. & Law, D. Digital Libraries. Aldershot, Ashgate
- LOCKSS(2009) <http://www.lockss.org/lockss/Home> (viewed 5 June 2010)
- McLeod, R., Wheatley, P. and Ayris, P. (2006) Lifecycle information for e-literature: full report from the LIFE project. Research report. LIFE Project, London, UK.
- NARCIS (2009) <http://www.narcis.info/index/Language/EN/> (viewed 5 June 2010)
- Nicholas, D. (2009) What is beyond books and journals? Pointers from CIBER's Virtual Scholar programme. Third Bloomsbury Conference on e-publishing and e-publications, 25 & 26 June 2009. Beyond Books and Journals. Programme available at <http://www.ucl.ac.uk/infostudies/e-publishing/e-publishing2009/> (viewed 5 June 2010)
- Openwetware (2009) http://openwetware.org/wiki/Main_Page (viewed 5 June 2010)
- Oxford (2009) <http://www.odl.ox.ac.uk/principles.htm> (viewed 5 June 2010)
- Roosendaal, H., Huibers, T., Geurts, P. and van der Vet, P. (2003) Changes in the value chain of scientific information: economic consequences for academic institutions. *Online Information Review* Vol. 27 pp. 120 - 128



DIGITAL LIBRARIES IN HIGHER EDUCATION

Tenopir (2009) The study is being conducted for Elsevier and will be published once results are finalised. The grant to fund this is noted at <http://web.utk.edu/~tenopir/research/grants.html> (viewed 5 June 2010)

Tenopir, C, King, D., Edwards, S., Wu, L.(2009) "Electronic Journals and Changes in Scholarly Article Seeking and Reading Patterns," *Aslib Proceedings: New Information Perspectives*, 61: 5-32.