
This version is available at https://strathprints.strath.ac.uk/8129/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. 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 (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk

The Strathprints institutional repository (https://strathprints.strath.ac.uk) is a digital archive of University of Strathclyde research outputs. It has been developed to disseminate open access research outputs, expose data about those outputs, and enable the management and persistent access to Strathclyde's intellectual output.
Digital library economics: aspects and prospects

Derek Law

Introduction

Although there is a substantial literature on digital libraries there is a huge gap when we start to look at their economics. A decade ago, Mike Lesk (1997) judged that ‘economics is definitely not a solved problem’. By 2004 he would write that it was ‘emphatically not a solved problem’ (Lesk, 2004). Writing in 2004 of The European Library (TEL) Project, Collier noted that:

... in 2000 it was very difficult to find published or public material on the business planning behind digital libraries. Published research work hitherto had focused almost entirely on technical, design, copyright and, more recently, human and social aspects. In comparison to the vast literature on these subjects, accounts of the business case behind digital libraries seemed to be rare or incidental to other issues. (Collier, 2004)

Charles Bailey’s bibliography (Bailey, 2008) runs to almost 300 pages but only nine, listing fewer than one hundred articles, are on economic issues. In practice most of these are about pricing models for journals, or electronic publishing, generally on the economics of information or scholarly communication or on consortial purchasing. The section on digital libraries has three times as many articles, but they focus on individual case studies or general overviews. Economic issues are notably absent. Halliday and Oppenheim (1999) undertook one of the first British studies trying to develop economic models in 1999 – ancient history in terms of digital
libraries – but that study focused mainly on the electronic journal. In the last five years there has continued to be a great deal of work on the economics of journal publishing and a growing literature on the economics of preservation and the economics of content creation. Further work has appeared on business cases for digital libraries or on particular services such as TEL or Europeana. But in truth, and almost without exception, we manage hybrid libraries and so the true economic cost of wholly digital libraries has not been fully identified, far less explored. This chapter then considers the aspects which have to be considered in the economics of digital libraries, areas where work has been done and areas where prospective work is required.

Business plans

As Collier ruefully notes ‘... many digital library projects are started without much attention to the business issues’ (Collier, 2004). As part of the planning for TEL (which has later extended into the Europeana Project) he conducted an extensive literature review which discovered a very thin base of published literature (Collier, 2005). That literature review remains a basic starting point for any exploration of digital library economics.

Since then, a few important documents have appeared. One manual (Bishoff and Allen, 2004) provides a useful template for all heritage organisations planning digital collections and helpfully identifies such items as indirect costs which should be included in any realistic economic model. However, this is a higher level document which, for example, quite properly proposes that, as part of the plan, staffing needs are addressed, but without specifying what these needs are. The Encyclopaedia of Library and Information Science (West and Strongrem, in Drake, 2003) offers the current state of professional thinking in a series of straightforward assertions about costs. Although no evidence base appears to support these, they probably do have elements of truth. It focuses on reducing support staff, cutting out departmental libraries and a reduced need for new space, for binding, for stationery, for reference staffing as footfall is reduced, and so on. There is, however, no real discussion of which costs will be modified or increased rather than lost. For example, it seems equally likely that some costs will move down the information chain to the user, either through their ownership and use of IT equipment or through the purchase or leasing of information using the
painless and plentiful micro-payment systems which exist. Costs can also be redistributed. One area of recent polemic debate has been the cost of setting up institutional repositories. The costs of such repositories will, of course, vary depending on the type of software and hardware used and the cost of research. However, some of those costs can be absorbed elsewhere in the system – for example, it has been argued that cost savings from the decrease in interlibrary loan traffic and the associated transportation can be transferred to creating a repository. These existing resources can be used to organise and run the repository (Ayris, 2005).

Some very elementary work has been done on attempting to estimate the economic value of libraries and information. Lesk (2004) has cited work undertaken in the United States and Canada which tries to monetise library activity. For example, the Minnesota Department of Transportation Library showed substantial cost savings in having a library as opposed to not (Baldwin, 2004) when some 4,500 information transactions were examined. The British Library (2004) has used contingent valuation to show that it creates value at the rate of 4.4 times the level of its public funding. Kelly et al. (2008) estimate that the value of knowledge transfer from the Scottish university libraries to the external community, as measured by library visits made by external users, is £7.1 million. This is a fruitful area for further study, although there must always be a suspicion that at least in the public sector the proper provision of and support for a good library is as much a statement of values and aspiration as a hard-headed economic decision. The financial benefit then simply becomes an argument to justify the prior decision.

Lesk (in Andrews and Law, 2004) has made some effort to understand library costs. A university will typically spend 3 per cent of its budget on running costs for the library. Half of that will go on staffing, one-third on the book fund and the rest on computing services and general supplies. He then looks at models for generating the revenue which will meet the cost of such libraries and is quite pessimistic that the money will be found. The TEL Project finished up taking a quite different approach and rested its business case on three important criteria which influence how we are to consider the economic aspects (Collier, 2004):

- digitised content (as distinct from licensed born-digital content) was likely to be important;
- TEL should concentrate on material which is not readily available elsewhere; and
- material should as far as possible be free at [the] point of use.
Note, however, that all three of these important principles focus on content and collection and not on services. Nevertheless there is a clear assertion that the library will be paid for from public and/or institutional funds. In this economic model, unlike Lesk’s approach, the business case will rest on providing value for money rather than on considering how income can be raised.

**Library costs**

We know quite a lot about existing library costs – at least about direct costs – through the long time series of the Society of College, National and University Libraries (SCONUL) statistics. But it is important to note that these simply do not cover indirect costs such as estate and heating and lighting, security and maintenance which are all increasing costs to the institution and costs which may look quite different for digital libraries. Nor has there been any significant analysis of these figures. There has also been some recent useful work on American public libraries, which offer toolkits for measuring value for money and cost benefit (Elliott et al., 2007). However, this does tend to focus on the physical assets of the library, which, being tangible, are easier to measure. Because it will be impossible to make sensible judgements about the cost of digital libraries until we have a sensible understanding of the cost base of our existing libraries, more work does still need to be done here. Perhaps the first real understanding of this need for comparative analysis can be seen in Montgomery and King’s (2002) work comparing the cost of print and electronic journal collections as Drexel University moved from one to the other. They were quite clear of the huge methodological difficulties they were tackling and stated with a becoming modesty that ‘this study should be viewed as a single first step to address an issue of critical importance to academic libraries.’

**The components of the digital library**

If we are to have a clear understanding of the economics of the digital library we first need to define it. There is no helpful agreed definition. Indeed:

To confuse matters, the job of libraries is expanding. Once upon a time, library patrons were always physically present in the building
where the books were held. Now libraries provide remote access across the campus and across the world. They do fax delivery, electronic displays, and other methods of information access. They provide sound recordings, videos, multimedia, and computer files. They have a tradition of providing access to books for no charge. Should that tradition also apply to these new services and these new users? Some of the new services appear to be filling needs previously filled by bookstores, journal subscriptions, and other paid-for activities. Some might be quite expensive to continue operating, for example, levels of computer support adequate to deal with much modern software. Does that mean that libraries should charge for these services, and, if so, on what basis? (Lesk, 2004)

The concept and initial experiments in digital libraries precede Web 2.0 by some years, but it is interesting to contrast how activities in the Web 2.0 world have supplanted the library and in many cases appear to have removed the need to interact with the traditional library (Law, 2007). It is then possible to compare how a digital library might be designed which in turn responds to Web 2.0 (see Tables 4.1 and 4.2).

This first attempt at defining the components of a digital library will, of course, need to be refined and expanded but it does begin to provide a template which will allow us to define areas of activity which can then be costed. As was always the case, the balance and cost of these various

<table>
<thead>
<tr>
<th>Traditional library activities</th>
<th>Web 2.0 activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cataloguing</td>
<td>Automated metadata, del.icio.us, zotero</td>
</tr>
<tr>
<td>Classification</td>
<td>Folksonomies, the semantic web</td>
</tr>
<tr>
<td>Acquisitions</td>
<td>e-Bay, PayPal, Amazon and Abebooks</td>
</tr>
<tr>
<td>Reference</td>
<td>Yahoo Answers, Wikipedia</td>
</tr>
<tr>
<td>Preservation</td>
<td>Digital archives and repositories</td>
</tr>
<tr>
<td>User instruction</td>
<td>Chatrooms</td>
</tr>
<tr>
<td>Working space</td>
<td>Bedroom, Starbucks with a laptop</td>
</tr>
<tr>
<td>Collections</td>
<td>Youtube, Flickr, Institutional Repositories, Open Access</td>
</tr>
<tr>
<td>Professional judgement</td>
<td>The wisdom of crowds</td>
</tr>
</tbody>
</table>
activities will vary from library to library, but as the evidence base grows it should be possible to benchmark against comparable organisations. It would clearly be of benefit if an organisation such as SCONUL took on the task of beginning to define and collect statistics and costs relating to the components of the digital library.

### Digital library collections

One significant feature for digital libraries which massively extends their power and range is the value-add which can be provided from making new linkages. The Aspect project was set up to create a digital archive of the ephemera – leaflets, flyers, postcards, newsletters – produced by candidates and political parties for the first Scottish parliamentary election in May 1999. The archive is based on the collection of election ephemera held by the Andersonian Library at the University of

<table>
<thead>
<tr>
<th>Traditional library activities</th>
<th>Web 2.0 activities</th>
<th>Library 2.0 world</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cataloguing</td>
<td>Automated metadata, del.icio.us, zotero</td>
<td>Metadata</td>
</tr>
<tr>
<td>Classification</td>
<td>Folksonomies, the semantic web</td>
<td>Locally provided and relevant folksonomy</td>
</tr>
<tr>
<td>Acquisitions</td>
<td>e-Bay, PayPal, Amazon, Abebooks</td>
<td>E-archives, e-data and quality assurance, ingest</td>
</tr>
<tr>
<td>Reference</td>
<td>Yahoo Answers, Wikipedia</td>
<td>Branded links to trusted resources, trust metrics</td>
</tr>
<tr>
<td>Preservation</td>
<td>Digital archives and repositories</td>
<td>Institutional repository</td>
</tr>
<tr>
<td>User instruction</td>
<td>Chatrooms</td>
<td>Moderate chatroom</td>
</tr>
<tr>
<td>Working space</td>
<td>Bedroom, Starbucks with a laptop</td>
<td>Wired campus and 24-hour workspace</td>
</tr>
<tr>
<td>Collections</td>
<td>Youtube, Flickr, Institutional Repositories, Open Access</td>
<td>Aggregation of unique content with other libraries</td>
</tr>
<tr>
<td>Professional judgement</td>
<td>The wisdom of crowds</td>
<td>Teaching retrieval skills</td>
</tr>
</tbody>
</table>
Strathclyde, which is acknowledged to be an important and unique record of a key event in Scottish history. The collection has been expanded through subsequent elections and makes links between constituencies, parties and individuals which hitherto would have been research projects in their own right. The same can be seen on a much more massive scale with the Voyages Project which has created the *Trans-Atlantic Slave Trade: A Revised and Enlarged Database, 1500–1867,* a huge piece of scholarly collaboration bringing together records from libraries on several continents.

Although it has been argued that what we have done so far is to create cabinets of curiosities rather than a coherent collection development philosophy for digital libraries, there can be no doubt that such examples of adding value signal one of the key benefits of such libraries. Again this is very difficult to measure, but in both the examples cited it would be possible to measure the changes in usage between phase 1 (digitisation) and phase 2 (making links which add value). Increased use will demonstrate that there is benefit from adding such value; quantifying that benefit is a much more daunting challenge!

So a key benefit of digital library collections will be the value we add to the products we acquire in order to make them more useful. A sufficient argument is that we need to do this in order to make the collection more relevant to the local community of which the parent institution consists. But there can be a much wider value. Lesk (2004) questions this, using the example of one specific data service set up by the University of Strathclyde. ‘BUBL has only 15% of its users from inside the United Kingdom. Many university administrators ask why their money is being used to pay for services when most of the people using those services are not from that university’ (Lesk, 2004). Yet libraries have always varied enormously in size, content and quality. Equally they have always seen themselves as part of a global system. Creating collections of local relevance but making them available globally is in the spirit of this tradition and in turn guarantees that the local users have access to a rich and global set of resources. It is not in principle different from the global inter-lending system which is one of the unacknowledged triumphs of the library profession.

**Digital content**

It is easy to forget just how recently the concept of digital libraries became current, perhaps a decade ago. What we have now is the early
modern Internet and what we are creating is the equivalent in timescale of incunabula after the printing revolution. We are in the midst of a frantic rush to transfer the huge wealth of analogue culture to the digital domain. Mass digitisation through such huge projects as the Google Books Library Project, which will capture millions of books in digital format, will inevitably transform how libraries are used and perceived, and yet it is still barely understood how the media of the past ought to be preserved, presented and interconnected for the future.

One of the first attempts to look at life-cycle costing comes from the substantial Life Cycle Information for E-Literature (LIFE) Project based at University College London (Watson, 2005). It aims at a comprehensive analysis of activities related to the management of content from selection, through licensing and acquisition to ingest, metadata creation, adding links, access, user support, storage costs and preservation. This provides at least the basis for a very robust methodology to define digital library costs. The project makes some quite specific estimates of costs. LIFE established that in the first year of a digital asset’s existence:

- the life-cycle cost for a hand-held e-monograph is £19;
- the life-cycle cost for a hand-held serial is £19;
- the life-cycle cost for a non-hand-held e-monograph is £15;
- the life-cycle cost for a non-hand-held e-serial is £22;
- the life-cycle cost for a new website is £21;
- the life-cycle cost for an e-journal is £206.

LIFE further predicts that in the tenth year of the same digital asset’s existence:

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

In some ways perhaps the most interesting feature of these very precise numbers is our inability to present similar figures for traditionally published materials. It is also important to remember 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. But these costs are just as real for libraries.

But there is another view of how library content and hence costs will develop in digital libraries. Often with great reluctance, librarians have become used to the concept, certainly for journals, of leasing rather than buying content. A marked reluctance to part with the hard copy still slows the development of truly digital libraries. The view is perhaps also coloured by the stately development of legislation on electronic legal deposit which will guarantee some level of future-proofing. Research libraries have traditionally seen their role as one of collection building for the future as much as about current provision. Digital libraries offer a radically new prospect. There must be at least a suspicion that content will be much more radically divided into the owned and the leased. The owned will broadly represent the intellectual output of the organisation in the form of research data, course material, grey literature and reports produced by staff, and some kind of publications repository. Larger research libraries may also choose to collect the electronic equivalent of archives and build and add value to collections as described above. However, it seems certain that the bulk of the recurrent collections budget will go on the leasing not the acquisition of data. For many years, collections were built ‘just in case’ items were needed. Library and library supplier automation and the growth of the Internet allowed some libraries to move towards a ‘just-in-time’ approach which rapidly acquires items when needed. Systems have now further developed which allow both electronic provision of content and hugely sophisticated tools which measure usage rapidly and efficiently and on the fly. It is at least theoretically possible to customise this rented collection on a daily basis to meet the changing needs of users. We may expect much more time, money and effort to be spent on this customisation to meet current need and less and less on collection building.

Digital library staffing

There is a rather facile view among university managers that digital libraries must inevitably lead to major staff reductions. Indeed, one suspects that some of the enthusiasm of university managers for digital developments lies in this thinking. Of course, on the surface the removal
of the whole lending, shelving and binding process would produce savings in staff numbers, while a reduction in staff numbers would in turn allow a delayering of management with even more substantial savings. Indeed, the prospect of replacing an increasingly invisible (since the library is rarely visited by academics who work from the desktop) and sclerotic senior management from the payroll would allow younger and nimble librarians to move the service forward. The recent JISC report on the eLib programme noted of librarians a general view that ‘theirs is a highly conservative and risk-averse profession’ (Duke and Jordan, 2006). And yet it continues: ‘Nevertheless, they have taken on board a huge raft of changes in the ways in which they deliver services. Indeed, they have taken almost complete ownership of the changes which have occurred, showing a degree of imagination and breadth of vision which is striking.’ The same view comes from Ayers:

The real heroes of the digital revolution in higher education are librarians; they are the people who have seen the farthest, done the most, accepted the hardest challenges, and demonstrated most clearly the benefits of digital information. In the process, they have turned their own field upside down and have revolutionized their own professional training. It is a testimony to their success that we take their achievement for granted. (Ayers and Grisham, 2003)

What we have seen and what we must expect to continue then is a re-skilling of the existing workforce rather than its wholesale replacement with a new breed of what Prensky (2001) calls ‘digital natives’.

And the pace of change shows no signs of abating. Prensky (2001) considers that the nature of content will change fundamentally:

It seems to me that after the digital “singularity” there are now two kinds of content: ‘Legacy’ content (to borrow the computer term for old systems) and ‘Future’ content. ‘Legacy’ content includes reading, writing, arithmetic, logical thinking, understanding the writings and ideas of the past, etc. – all of our ‘traditional’ curriculum. It is of course still important, but it is from a different era. Some of it (such as logical thinking) will continue to be important, but some (perhaps like Euclidean geometry) will become less so, as did Latin and Greek. ‘Future’ content is to a large extent, not surprisingly, digital and technological. But while it includes software, hardware, robotics, nanotechnology, genomics, etc. it also includes the ethics, politics, sociology, languages and other things that go with them.
This rather chilling view of the sort of content we will have to curate implies a huge change in staffing skills as well as structures. Cannell (2007) is one of the few to have thought about the changing skill-set. She tabulates it as:

- Acquisition and cataloguing becomes ingest and metadata
- Static service desks becomes helpdesk and rovers
- Collection development becomes procurement
- Subject librarian becomes support of learning and research, kitemarking
- Archival skills peripheral becomes archival skills central
- Systems librarians become software and Web 2.0 developers
- Preservation of print becomes digital preservation and data curation

It would seem likely from this that while staff numbers may decrease, the broader and richer skill set means that staff costs will not. A huge amount of work is needed to think through and cost the sort of balance of staff required for digital libraries and then to compare that with existing costs.

**User support, kitemarking and trust metrics**

Geoffrey Bilder (2005) has written extensively on the subject of trust metrics. The classic trust marker was and is the kitemark used for all sorts of products in the UK. In libraries possibly the most significant markers were a publisher’s reputation for books and journal titles for serials. These begin to disappear or at least erode on the Web. Another implicit metric was the very fact that the book was in the library. This implied a filtering and selection process by academics and librarians which meant the book was relevant to and fit to be in the collection. Cannell rightly identifies the importance of this filtering and kitemarking in the digital library as a task for subject librarians. This validating of web resources will add value and minimise the time required by searchers and researchers. By, in effect, guaranteeing the quality of the information source the library will add another very complex piece to the economic equation.
The estate

One paradox of the arrival of the digital library is the literally hundreds of millions of pounds spent in UK higher education alone to create new buildings. A plethora of concepts from traditional libraries through learning commons to idea stores is being built. As a rule of thumb – which needs to be tested – it seems not unreasonable to estimate the typical annualised cost (capital cost ÷ building life in years) as being in the millions and broadly equivalent to the size of the book fund. However, we also need to factor in maintenance and utility costs. Libraries are expensive to light and heat and are noticeably profligate at a time when utility costs are soaring and institutions are increasingly conscious of green agendas and sustainability. Although there is a general wish to safeguard the library as a place, whether as a learning or simply a social space, there can be little doubt that the wholesale disposal of paper and in particular science, technology, engineering and medicine (STEM) journals which are available electronically offers the prospect of significantly reduced space footprints. In turn this will reduce both capital and utility costs. Of course increased use of computers in the library may increase utility costs, but these do need to be quantified.

Administrative costs and income

The serious question is whether there is enough financial support for the information on the Web. Commercially supplied information will presumably exist while it generates income. Ideally it will then migrate to legal deposit libraries. But much (most?) of the information on the Web is not commercial and it is important that we develop a stable infrastructure which ensures the long-term future of non-commercial information. Since this is broadly the equivalent of special collections and archives in traditional libraries, they have a key role in developing the structures which will ensure such permanence, irrespective of financial reward.

There are of course revenue streams which libraries could improve or develop. A prevalent working assumption is that library photocopying income is in steady decline as students turn to printing pages from the Web. If so, that pot of money is available for the purchase of information. Another working assumption is that the normally low level of library fines allows at least some students in effect to rent textbooks. This money
is also then available for the leasing of electronic information. Although figures are very patchy and institution-specific, IT ownership is now the norm for students with laptops prevalent and mobile phones ubiquitous. The increasingly common provision of desktop machines in libraries with expensive replacement cycles will surely and quickly be replaced by small portable devices owned or rented by the reader. Indeed as battery life improves it may not even be necessary to provide access to power supplies. Of course IT/information staffing support costs may rise as a result but at least some institutions offset this cost by selling a preferred device to students, sometimes with course content loaded, and with a margin for support costs built in to the price. Support and administrative costs are an extremely difficult area to quantify, but they undoubtedly exist. Many will be the same as in a traditional library, but almost by definition IT support in the broadest sense will be much more significant while software and data licenses will also be a new or at least much larger component.

**Conclusion**

In the absence of any significant body of work on the economics of digital libraries or how to compare them with the economics of traditional libraries, this chapter has attempted to set out the issues and activities which require further exploration. When everything – or at least enough – is available on the Web, why will institutions bother to have libraries at all? We now know a great deal about the price of electronic content, but almost nothing about the total cost of its ownership or how that compares with traditional bricks and mortar. We know very little of what should be measured, far less how to measure it; but then it is only in very recent years that the concept of looking at the economic impact of traditional libraries has been explored, so this is perhaps unsurprising. Having determined that for practical as well as philosophical reasons there is an economic benefit in retaining libraries, we shall still need to explore how they are to be funded. A study of Digital Library Federation members in 2001 (Greenstein and Thorin, 2002) concluded that core funding was required in some measure for support of the digital library, not least as it moves from projects into the daily mainstream of library life. There was clear evidence of library funds not dedicated to digital content per se being used. Digital library activities are being funded by money not earmarked for that purpose.
It seems safe to assume that, as with conventional libraries, there will be a huge range of levels of support. The real question we will have to answer is how far we provide value for money and how far we add value compared with what will increasingly be seen as the alternative of a web-dependent but library-free world.

Notes
2. http://gdl.cdlr.strath.ac.uk/aspect/

References
Digital library economics: aspects and prospects


