PARLOUR GAMES: THE REAL NATURE OF THE INTERNET

Derek Law

The future of the information society lies not in the "howling wastes of the Internet" but in the cost-effective library, that concentration of resources and skills. As network charges at site level are introduced and originators of information realise its value as a commodity, Metropolitan Area Networks offer the librarian the opportunity to develop the role of intermediary in an economical hybrid library, a role which will require a redefinition of services, greater emphasis on use support with proficiency in information skilling and training. In order to take the opportunity librarians should cease being passive spectators of the changing information scene and become the SAS of the information world.

Derek Law is Director of Information Services and Systems, Kings College London, Strand, London WC2R 2LS E-mail: derek.law@kcl.ac.uk

Everyone knows, albeit usually slightly misquotes, the famous lines,

> "Will you walk into my parlour? " said a spider to a fly:
> "Tis the prettiest little parlour that ever you did spy."[1]

But hardly anyone knows their nineteenth century author, novelist and translator Mary Howitt. Let us then take a look at the pretty parlour where the web exists and at some of the darker corners which lie behind the prettiness in order to test some of the assumptions already made about electronic libraries and to provoke some careful thought about what the future holds.

The failure of the Internet

The sudden mushroom growth of the Internet has caught most of us by surprise. Ten years ago the MARC User Group conference on Networking[2] has no index entry for it; five years ago the first UKOLN Conference on Networking and the Future of Libraries[3] has thirteen index entries, with over half of them in two papers. The published proceedings of the second UKOLN conference in 1995[4] has an index entry with twenty-one sub-headings. In the United States higher education is already building what it calls Internet II. And yet I want to argue that the Internet has already failed. It is clearly a technical wonder, but its philosophy is static and superseded and it will come to be seen as a milestone on the road to somewhere else.

The introduction of JANET charges will come in twelve months. For most institutions these will not be large to begin with, but as higher and higher bandwidth is sought, they will soon increase. This will almost certainly be done through charging rather than an increased topslice to fund JANET. Institutions will then almost inevitably ask who requires the bandwidth and for what, or else will pass on the charge to departments - who will ask exactly the same question.

There is an assumption that the Internet's reach is universal. While this is true it disguises the complete lack of thought given to good network topology. All network users know that the United States ceases to exist in the European afternoon and that a 404 message or a DNS failure really only means failure to access not denial of existence. Any decent university needs global not just North American links, but no visible design effort is going on to ensure that resources are mirrored and cached at network nodes where access can be guaranteed.

Present structures allow unrestricted access to irrelevant content, the acquisition of what is euphemistically defined as flesh-toned images. No organisation can sensibly offer an open-ended blank cheque book approach to information provision. This is entirely to do with finance rather than censorship. Yet, for rather woolly liberal notions, we tend to offer unrestricted access to the Internet in a way that we do not dream of doing to printed information.

Some recent quotations from the *Times Higher Education Supplement* give a firm impression of the issues facing the Joint Information Systems Committee of the Funding Councils, as it continues to invest in "one of the unsung successes of the British higher education system ...the academic electronic highway JANET and its broadband successor SuperJANET". But while the physical JANET network remains the best bargain for decades, a policy for content provision has been slower to emerge. If it is true that "the rewards are likely to be dramatically greater than the costs, as people become familiar with new ways of gaining information", it would be a great mistake to assume that the great and growing volumes of information on the Internet have equal - or indeed any - value.

"The howling wastes of the Internet" is a phrase I love. It very aptly describes not only the featureless landscape, but the failure to provide any signposts. Now of course we are working on that. Most commentators agree that perhaps 80% of Internet content is rubbish, so what do organisations such as OCLC do? They catalogue it: a classic librarian's response. Resource discovery systems from ADAM to

EEVL have more sane approaches to collecting data on a subset of purely relevant information, but it is as yet unclear how widely applicable their selected set will be.

Of course very basic search tools do exist, the web-crawlers. They are easy to use but represent the worst features of the dumbing of the Internet. We have created a new class of user, the satisfied inept, those who think they have the whole answer when they do not. They have mastered the technology, but have failed to recognise that information management poses a more complicated set of problems. The computer was once defined as a very fast idiot - and that remains the case.

So the Internet brings many problems in its wake and may not be the Holy Grail users have sought.

The success of the intranet?

As we search for new models for network management the library presents itself as an interesting analogue to the intranet. A library represents an attempt to concentrate the information resources and skills necessary to the organisation. No university would give staff and students an open account at any bookshop in the world - including the top shelf of the local newsagent's shop. Typically material cited as relevant to courses is concentrated in one building as is much of the material required by academic staff. Material acquired from outside is provided by intermediaries who do this in a cost efficient way.

In the same way we can imagine an intranet where the institution spends on servers as an alternative to bandwidth. In this model there is, as with books in a library, a filtering of acquisitions to ensure relevance and competitive pricing. It has been received wisdom for a decade that most libraries must move from a holdings to an access strategy. It is then a neat paradox that improved communications, poor network topology and the introduction of charges at site level in higher education may reverse the economics and force us to consider whether holdings are not to be preferred to access in many cases. It may be found that mirroring and caching of data, at least at Metropolitan Area Network (MAN)

level, is the most economic model. Thus the intranet reflects the concept of the library with a collection of relevant material made available by professional intermediaries in the most efficient and economic way. In addition, the material is so organised that it is preserved for the future, while all the issues to do with rights of access and obligations of ownership are properly managed by these same professional intermediaries.

If the principle commodity in the information society is intellectual property, we may expect all organisations to be much more aggressive over rights management. As a first step, our organisations will have to pay much more attention to the information created within the organisation. I would venture that no university and possibly few other organisations have an accurate and comprehensive view of the data resources created and managed by its staff - and certainly have no policy on its preservation and conservation, rights of access and obligations of ownership.

What makes this model both more attractive and more attainable is the arrival of the Metropolitan Area Network. It becomes possible for groups of organisations and not just those in higher education to co-operate in resource acquisition and provision. This has obvious cost-saving potential. Thus far the MANs appear to have been dominated by technical considerations, although there are welcome signs of emerging thinking, particularly on cross-sectoral co-operation. Greenstein has commented[5] on the democratic impulse which has inspired the Funding Councils to make more content available on the network and to remove Naylor's "tyranny of distance"[6], allowing researchers to work collaboratively. The new government has espoused a philosophy of regionalism. This was prefigured in the Anderson Report and the Funding Council's decision to change the use of DevR funding to allow weak research departments to work with the strong, further reinforces this thrust away from competition and towards collaboration.

We may expect to see the emergence of, in an expressive phrase, server farms. Specialised servers can be shared or networked, each covering a specialised area but offering a

comforting level of redundancy. For some activities, such as mail, each organisation may require its own server, for others, such as images, sharing may be both more appropriate and cheaper.

The hybrid library

JISC Circular 3/97 has usefully introduced the concept of the hybrid library, which encapsulates the notion that major institutions will have to work in an environment where access will have to be provided to a wider than ever range of material types. Further, there is an ambition to ensure that there is integrated access to this range of resources, rather than serial searching of differently structured databases from different curatorial traditions. Although fairly evident it is perhaps worth reminding ourselves of the range of materials involved.

Archives

The Non Formula Funding for the Humanities provided by the Funding Councils has provided a long overdue boost to this area of basic research material. The range and location of such material has often been one of the best kept secrets of higher education and this programme will begin to make these riches more available

Paper-based collections

A useful development within the hybrid library will be the integration of catalogues showing the complete range of what the library can provide. At its simplest level this means linking journal title records to abstracts and indexes, but also opens the possibility of linkages to such areas as conference papers. Even in the paper-based collections some substitution has gone on over the years, most notably with microform, although that tends to be integrated and conventionally recorded in the library catalogue.

Images

Higher education produces a vast range of images each year ranging from dental and x-ray to fine art. It also uses a vast array of images, both moving and still. Perhaps the best recent example of this is the Visible Human Project, a

huge bank of images created by the National Institutes of Health in the United States and mirrored in the UK at the University of Glasgow.

Sound

Again there is a huge variety here from the sound of heart murmurs in medical education to music itself via public service broadcasting. Other material can give a profound sense of time and place, whether Martin Luther King's electrifying "I have a dream" or Neville Chamberlain declaring that a state of war exists between Britain and Germany.

E-journals and e-books

The Pilot Site License Initiative has done more to make electronic journals generally available than perhaps any other initiative. Significant numbers of journals are now available electronically although the vast bulk of them remain copies of the printed version, rather than innovative new material which takes advantage of the possibilities of the network. Electronic books have perhaps been slower to develop, although many texts are available on the network from sources such as the Oxford Text Archive.

Grey literature and pre-prints

These are a staple of research in some disciplines such as economics. If the revolutionary work in this area was done at Ginsparg's famous Los Alamos archive, the UK has been quick to follow. JISC has funded projects covering economics working papers and a Cognitive Sciences pre-print archive at Southampton under the control of Steven Harnad. Nor should one forget the faithful and longstanding efforts of the British Library to catalogue grey literature.

CD-ROM

Most higher education libraries will now have a large number of CDs available and the majority have made some effort to network these with greater or lesser success. There is a growing body who use such products as Ovid, having recognised that there are prudent limits to what can be achieved with this technology.

Datasets

Perhaps the defining new resources, the "huge leap forward"[7], which gave an undoubted impetus to the electronic library in the UK are data services such as BIDS. But datasets now cover a much wider range of materials than secondary bibliographic resources. Satellite data, chemical data and basic research data are all now readily available to the academic community.

Resource discovery services

This is the last element to be mentioned in terms of the hybrid library. JISC has funded a large number of discipline based resource discovery services. Others are springing up whether nationally or internationally. These are clearly both popular and effective, but it may be that over time they come to be seen as a nationally or internationally provided core, to which local information and value is added either institutionally or regionally. Almost every discipline has a local dimension, which has to be accounted for somehow. For the moment, this possibility must again remain speculation. Even this list ignores a range of other resources ranging from software to video and locally created resource packs.

The emerging experience

The first evidence is beginning to emerge on the use of electronic and hybrid libraries. The experience at Tilburg University is perhaps best recorded.[8] Crudely the students use the library more often and spend more time there, while the faculty members use it less. This leads in turn to a requirement for a much greater volume of general technical computing support within the library, coupled with a need for much more targeted specialist support for research staff at the desktop. Some of this need for increased technical support can also be seen in the UK Pilot Site Licence Initiative[9] for journals, where the provision of electronic versions introduces new issues for libraries, such as how to support Adobe Acrobat.

Converged services also tend to be responsible for what, in a clear but infelicitous phrase, are described as mission critical

activities. Typically these will be electronic
mail, CWISs and webservers rather than our
beloved OPACs. It is then perhaps surprising
how little resource and resilience we provide for
such services. The new environment will not
just make resource demands, it should also
force us to consider afresh how we attribute
resource in support of institutional mission.

In United Kingdom higher education the
managerial solution to these emerging issues
has been to create some kind of academic
services or learning resources directorate, under
a single director and with library and
computing centre at its core. Half the
universities in the UK now follow this model,
with many still changing as the opportunity
presents itself. Very few have consciously
chosen not to follow this route when the
opportunity has appeared. Yet this is (with
isolated exceptions) a peculiarly British
phenomenon. It is often blamed on the Follett
Report, except that much of the convergence
predates Follett and the report did not
recommend it. What the Follett report did do,
was to cause all institutions to consider
information strategies and this may be the
underlying cause of change. It will be
interesting to see whether the fashion survives
the first generation of postholders (this seems
likely) and whether the fashion will spread to
other countries.

The information intermediary

If we are then to play parlour games we should
define the strengths we have and play to them,
and define the skills we require and train for
them.

There is a clear and growing need for
enhanced user support. This will be more
varied than ever, ranging from software and
printer problems to content and HTML
formatting problems. Furthermore, the user will
expect library and computing support to be
both integrated and available when and where
the user needs it. Coupled with this users will,
wittingly or unwittingly, require increased
training, again at the time and place of their
choice. The huge growth in the student body
and in the range of information types and access
routes to them, required by even the most junior

user, make a daunting combination. Training,
information skilling and even curricular
teaching may well become the dominant
requirement of librarians over the coming years.

We are all led to believe that information is
increasingly a commodity. Commodities are
traded and information arbitrage, the
acquisition of data at the best price, is an
intriguing possibility. Certainly, for information
available from multiple suppliers we may
expect different prices from different suppliers
offering different services or the same service at
different times of day or in different parts of the
world. Conceptually at least, we could buy data
cheap from an Australian supplier if used only
in UK daylight hours, when simultaneous
Australian overnight use is minimal, or like
industrial interruptable gas supplies, we could
have little used data available only when traffic
loads within the UK are light. BIDS usage
figures have shown that there are sad people
who use the network even on Christmas day
and New Year's Day.

Perhaps we should also look at some of the
commercial practices of telecommunications
companies. Is the virtual reference desk a
possibility for example. If directory enquiries
for London can be handled from Inverness, how
much helpdesk and front line support could be
delivered electronically or by telephone from
homeworkers or by a local consortium? Does
every library in London University have to staff
a support desk from 7pm to 10pm?

The Metropolitan Area Network offers one of
the best opportunities for this. It has produced
massive but almost unnoticed deregulation and
is an opportunity at present.largely going
begging or being seen in terms of linking
OPACs. Some real vision is needed to grasp
this huge opportunity for collaboration,
development and enhanced service provision. A
range of activities from resource discovery to
datasinks and server farms could be explored
on a MAN basis and indeed the Dearing Report
is widely expected to encourage this.

Parlour games

The content of the parlour games of the title has
largely been described above, but in closing we

can link the descriptions quite readily to traditional leisure pursuits.

Charades (or copyright clearance)

King's College offers over 1,200 course modules. Let us suppose that each has a fifty-item reading list which we want to make available electronically and that King's is of average size. For the system as a whole this is 60,000 copyright clearances times 100 universities, which make potentially about six million clearances a year - or more accurately, requires a system capable of delivering this in a fortnight. Although it is avowedly the start of the process, projects such as those at De Montfort or Loughborough working on rights clearance, make depressingly slow progress. Compare my six million with Project ACORN's clearance for digitisation of 158 journal articles with 51 publishers.[10] Why the sort of sampling which is acceptable for public lending right (PLR) cannot be used, I cannot imagine. Of course, PLR is aimed at authors. Only the most naïve and unsophisticated emperor watchers might suppose that publishers would not want to let authors know which works are being borrowed.

One other possibility sometimes touted in the Funding Councils is to set up a Copyright Licensing Agency for HE: an agency through which the universities sell time- or medium-limited rights to publishers. As universities become more alive to intellectual property rights and the way in which the Web has become a low entry publishing medium, they must consider protective action. As institutions look to position themselves in an information society the conclusion seems inescapable that it is the possession of intellectual property rights which will give stakeholders a uniqueness and value. It is inconceivable that universities, research bodies, grant awarding bodies and authors will continue to sign away their rights in the present cavalier fashion

Sardines (or collection building)

Almost by accident the UK may have begun to diverge in important ways from the United States and other parts of the world. Thus far the principle thrust of digital library thinking has

revolved around the scholarly journal, whether through digitisation of old content, simultaneous paper and electronic publication of new content, or replacement of the current journal style with a multimedia extravaganza. This thinking has been shared globally and has concentrated on the products and outcomes of research. However, in the UK, most of the expenditure has quietly gone into the acquisition, organisation and preservation of the raw material of research, ranging from satellite data to mediaeval manuscripts. A growing number of funding agencies are making it a condition of grant that research results must be offered to one of the national data centres for preservation. It can surely not be too long before institutions also recognise the value of having a firm policy in this area in order to protect the intellectual capital they create. Whether this divergence by the UK is more apparent than real may not become evident for some time, but forms an interesting speculation.

Hunt the Thimble (or resource discovery)

A great deal of effort is going into this area and both the Dublin Core and its successor the Warwick Framework show the UK to be much involved in the process. If resources on the Internet are to be made usable it is self-evidently important that they must be identified, recorded and made accessible. At present there is a jejeune assumption that more is there than in reality exists, and conversely, that a search undertaken by one of the major search engines removes the need to consider where relevant non-electronic resources may be identified and found.

Monopoly™ (or commercial suppliers)

This needs little explanation. The cost of research reported in academic journals and funded from the public purse is about one million dollars a minute. Having completed the research and written it up, the intellectual property rights are given away, particularly in the sciences, to a very small self-perpetuating oligarchy. Then universities buy it back. The system is quite capable of self-reformation but

as yet awaits the development of a large enough cadre of fanatics seeking such change.

The Grand Old Duke of York (or training needs and user support)

These are, I believe, the key issues that we need to address as a profession. A typical university now recruits several thousand students each year. All too often induction consists of marching them up to the top of the library then marching them down again. There is both scope and need for imaginative redefinition of how we deliver what users need, when and where they need it.

Pooh Sticks

This is the last game to mention and one we normally play but should now give up. As a passive spectator sport it suits many librarians temperamentally, but we can have little optimism about a philosophy which relies on seeing where the currents in the stream will take us, which will leave some trapped in backwaters, some beached on the shore, some moved at impossible speed and only a very fortunate few in calm water.

Let me finish then as I started, with Mary Howitt. She wrote some 110 books, in the main jointly with her husband William. The *Dictionary of National Biography* reports the *Times'* view of them "Nothing that either of them wrote will live, but they were so industrious, so disinterested, so amiable, so devoted to the work of spreading good and innocent literature, that their names ought not to disappear unmourned". So will librarians, those spreaders of good and innocent literature, be flies whose name disappears unmourned as they are replaced by cybrarians? Or will they be the spiders decribed by Miles Kington:

"Spiders are the SAS of nature, and will spend hours flying through the air on their ropes, prior to landing and subjecting some

hapless insect to savage interrogation. The question they usually ask is: 'Have you any last requests?'"[11]

References

1. The Spider and the Fly, a poem by Mary Howitt, 1834.

2. Horsnell, V. Mechanisms for cooperation. Aldershot, Gower, 1988.

3. Smith, John W.T. Networking and the future of libraries, London, Meckler, 1993.

4. Dempsey, L., Law, D. & Mowat, I. Networking and the future of libraries 2: managing the intellectual record. London, Library Association, 1995.

5. Greenstein, Dan. Lecture to Libraries and Information Commission given at King's College London, May 1997.

6. Naylor, Bernard. The convergence of the library and the computing service: the central issues *British Jnl of Academic Librarianship* 3 (1988) pp.172-187.

7. East, Harry et al. A Huge Leap Forward: a quantitative and qualitative examination of the development of access to database services by British Universities 1988-1994 (British Library R&D Report 6202) London, British Library, 1995.

8. Geleijnse, Hans. A strategy for information access *in* Dempsey, L., Law, D. & Mowat, I. Networking and the future of libraries 2: managing the intellectual record. London, Library Association, 1995.

9. Report on Phase 1 of the Evaluation of the UK Pilot Site Licence Initiative. Bristol, HEFCE, 1997.

10. Gadd, Elizabeth. Copyright Clearance for the Digital Library. *Serials* (10) 1997, pp27-31

11. Kington, Miles. Nature Made Ridiculously Simple (1983).