

And What of Intellectual Landscapes in the Future?

J.P. McCarthy

Boole Library, University College Cork, Ireland

JP.MCCARTHY@UCC.IE

Abstract

The world of librarianship has undergone a sea change in its understanding of itself and of its role as a contributor to scholarship, teaching and learning during the past twenty years. It now seems poised to facilitate the opening of new vistas on future knowledge access and interpretation. It has become a leading force in the evolution of new intellectual landscapes while at the same time becoming ever more conscious of its traditional custodial role in preserving the media of knowledge transmission. There are so many facets of change taking place, so much research, so many reports, so many scholars and so many commercial companies contributing to create the knowledge society; a knowledge economy. So much of the language and culture of the knowledge society is derived from the world of commerce. Its evolution seems to be more and more market driven. But what might happen if knowledge was no longer the focus of the marketplace? Would production cease? This paper speculates about such a future and the knowledge landscape which might emerge.

Keywords: Book, Intellectual landscapes, Grid, Information, Knowledge, Future, Literacy, Librarianship, Digital, Repositories.

Introduction

Project your mind to a paradigm in the year 2500. One where there is a universal corpus, in digital form, of all recorded knowledge. One where there is universal availability of such a corpus, though not universal access. Try to imagine how this paradigm might have come to be.

Crest of the Wave

According to a report produced by the School of Information Management (SIMS) at the University of California, Berkeley in the year 2003 (Lyman & Varian, 2003) the quantity of information being produced by world society can be represented in two categories i.e. physical media and information flows through electronic channels. The four physical media are print, film, magnetic and optical.

The four information flows are telephone, radio, TV and the Internet.

Material published as part of this journal, either on-line or in print, is copyrighted by Informing Science. Permission to make digital or paper copy of part or all of these works for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage AND that copies 1) bear this notice in full and 2) give the full citation on the first page. It is permissible to abstract these works so long as credit is given. To copy in all other cases or to republish or to post on a server or to redistribute to lists requires specific permission from the publisher at Publisher@InformingScience.org

The report estimated that the quantity of new information generated in 2002 was in the region of 5 exabytes through physical media and almost 18 exabytes through information flows.

The report defines 5 exabytes of information as being equivalent in size to the quantity of information con-

tained in 'half a million new libraries the size of the Library of Congress print collections'; if digitised. In 2002 ninety two percent of the information stored on physical media was stored on magnetic media, mostly in hard disks. The authors estimate that the storage of new information has been growing at a rate of 30% a year. This typology of information production and dissemination has evolved since prehistoric times; from the first use of artifacts for the transmission of information. Today we ride the crest of this wave. Where will it take us from here and how much control will we have over it?

In my imaginary scenario of 2500 the wave has subsided and a vast pool of information exists. Production has ceased. It is not structured, there is no architecture, no towers, no peaks and troughs, no physical intellectual landscape and no predefined navigation routes. But when required a net can be lowered onto the pool. It is a grid which can create navigation routes for miniature robots/webcrawlers. These can extract information from the pool and interact to define, refine, package and deliver the product to a seeker's specifications.

The Lines of Change

How might the elements described above have come about i.e. the pool, the net/grid, the robots, the mindset of the seeker? It has been said that technological revolutions are not the result of revolution but of many small independent developments reaching a critical mass; one where the past has slipped out of vogue almost imperceptibly. Let us look at a historical scenario for each of the elements above to see how what is happening today might have evolved into our vision of a future.

The Pool

In describing his vision of the World Wide Web Tim Berners-Lee (2000) speaks of a space where potentially anything can be linked to anything else. We see evidence of initiatives towards harmonisation in many areas of the information world today from metadata markup language standards to encoding levels for interoperability standards to facilitate long term migration and the consequent preservation, of cultural heritage stored in digital form. In traditional knowledge storage and dissemination, society created places where the encapsulation and transmission devices used to store such assets could be gathered together. These tangible devices have dictated the existence of a part of the built landscape and within them are those structured assemblages of the printed word that we refer to generically as 'books'. The physicality of handling such artefacts dictates the need for organisations to exist to control access and maintenance. The whole panoply of resources required to deliver access to the corpus of human knowledge, intangible information objects, represents a vast investment over time by society. How much of the total corpus has ever been used/exploited? For the sake of argument, if only 50% of what exists has ever been tapped then why is this the case? Is it because of the classification systems we use to label and arrange what is there? Is it because of the extent of cataloguing coverage? Is it because of dispersal across the global landscape? Across the libraries of the world we invest heavily in improving access to our collections to partly answer these lines of questioning while in parallel the Information Technology industry and information professionals generate an alternative world. Libraries are said to be in a state of transition which presumably means a state of shifting from being repositories of printed knowledge to being access points/control points for intangible knowledge assets, most or many of which are not owned, possessed in a traditional sense but simply subscribed to by licensed agreement. You will see much debate in the literature of Librarianship about the issue of hybrid libraries and their continued existence. You will also see forecasting which suggests that in time the shift will be total. Should that happen the waters of the eight information streams which now flow will become the pool.

The Net/Grid

Since the beginning of machine readable cataloguing the use of markup languages and frameworks for encoding cataloguing information into manipulable data elements has been underway. Libraries have used the MARC standard for this. Similar developments are underway with regard to the use of cataloguing standards for describing collections rather than individual items and for this process we use derivatives of the Dublin Core elements. As museums, art galleries, archives and libraries, begin to adopt the same data description standards the possibility of drawing from all of these independent records of cultural heritage through a single search statement opens a new vista for the knowledge seeker. From our OPAC catalogues we can now place hyperlinks to electronic resources as well as physical resources. Our catalogue entries rather than being just directional tools to physical locations for books and academic journals are now also windows to the intangible. A student can sit at his study desk in his accommodation unit and access all of his course material as well as a significant number of full text reading resources without ever visiting the campus library. The modern student is increasingly a network user, a cyberspace navigator. Initiatives on a European scale, in time perhaps on a world scale, will result in comprehensive full access through an international grid system to an already vast amount of knowledge stored in institutional and corporate computer systems. Therefore, the network of potential interlinking grows. The Grid and its routeways of ubiquitous persistent addresses grows exponentially to facilitate interlinking. But what of the traffic along the gridlines? What carriers will facilitate the student's quest?

The Robots

What of the software information scouts and their reporting habits? What of the emerging landscape elements of the superhighway – its hubs, gateways and portals, its access protocols and tariffs, its search engines, web-crawlers and spiders? Each of these categories of software tools influences the nature of the traffic on the Grid. They are a means of scoping for knowledge resources, a means of opening up access. In a sense they replace the physicality of the traditional knowledge seeker in a library. Are they worthy surrogates, and will their evolution influence the way we select and use knowledge? Could their evolution influence the cultural habits we have developed for accessing and exploiting knowledge/information? Might they create their own degree of obscurity as happens within the print collectors' world: as a result of design philosophies for organisation and arrangement or because of the limitations of the technology? Whatever the caveats may be, these robots may indeed control/define our future relationships with the Grid and the knowledge resources it overlays.

The logic given to these robots coordinates information at the time of discovery. In a sense this removes the need for a pre-coordinated architecture of knowledge as is the world of books. The architecture of the grid comes into being only when called. Such technology allows for its own ongoing evolution independent of the knowledge pool. Therefore the grid is lowered into place at the choice of the individual navigator. The architecture of the search is defined at the moment of the search, the architecture of the informational objects retrieved is defined at the moment of on-screen assembly. It is not predefined by the past. Witness subject cross-searching of electronic textbooks as an embryonic example of this.

The Seeker

Which brings me to the issue of the mindset of the information seeker. What mental attributes make a good researcher in the world of the printed book? Will this set of abilities and skills need to change to take full advantage of the opportunities provided by the digital world? Are librarians planning to deal with this and what developments are currently underway? There is a significant and growing body of literature on Information Skills Training and there are many initiatives and

standards being set internationally whereby librarians are either actively engaged in this type of work independently or as members of cross disciplinary curriculum design teams and teaching teams. Such activities give the Librarian a newly defined role as trainer, guide and advisor. Perhaps it is they who will ultimately define the seekers relationship with the pool; as they have always done within the print tradition. Perhaps it is they who will define and refine the usefulness and value of the pool for the seeker and thereby model the mindset of the future researcher.

The Nature of Technological and Social Change

In their seminal work *The Coming of the Book*, Febvre and Martin (1997) describe the origins of this information transmission and storage device and in so doing provide insight into the manner in which social need for the product emerged and developed. We are shown the industry of text production, its associated materials industry and skills input in the environs of the medieval universities and we get a sense of how market demand was facilitated by the new product. We further see how the product gradually spread from these nodes into so many areas of secular life; a slow infusion which has brought us to today's knowledge society concept. Along the way there have been many turbulent evolutions in the creation of knowledge and in gaining access to it. Examples reside in the intellectual history of the Age of Enlightenment, the Industrial Revolution, Victorian social reform movements, the Information Explosion of the 1950s. We also read of forecasts that did not materialise i.e. the paperless office; though perhaps it was only that the timeframe predicted for its realisation was too short.

What can we learn about possible futures from this history? Forecasts are seldom accurate. Core concepts of change that one-day seem naive can become commonplace overtime; meeting the challenge to the entrepreneur and venture capitalist. New forms of knowledge use and new concepts of value can be absorbed by society quickly as they may better facilitate current need than the older forms of technology. New perceptions of need can develop along with the tide of change and expectations can become commonplace standards. Therefore the technology of the printed word may succumb to such changes in the common social psyche; and to an extent has already done so as witnessed by the variety of media in use today. Digital technology brings new challenges as a new common storage and delivery base emerges i.e. the web along with newer configurations that may supersede it. So the lesson here is perhaps that knowledge currently stored in many formats is gradually merging into a single pool with an emerging overlay technology to exploit it, to integrate it, to co-ordinate it. To inter-relate it in ways never dreamed of before, to lift it from the constraints of the parietal media which have held it.

The Tools of Re-engineering the Print Block

Today we speak of electronically born documents and of those which are 'traditionally' created i.e. encapsulated into the world of print-technology. We think of e-born texts as those which have an existence which will never be part of the world of printed text unless an individual reader decides to take away a personal print copy. Spearheading initiatives in academe which encourage digital authoring and archiving are organisations like SPARC. The growth of electronic archives at many university sites is a physical manifestation of academe poisoning itself for change in one of its modes of communication i.e. the medium of journal publication. We see initiatives like JSTOR creating retrospective corpuses of journal titles while at the same time participating in initiatives to preserve a print archive of such titles. We encounter other initiatives such as Early English Books in Print, Eighteenth Century Collections Online capturing part of the corpus of printed literature in English in database form. We encounter microfilm archives of newspaper runs being converted to digital ones while at the same time promoting a heritage backup of both print copy and microfilm master. There are many more such initiatives across the traditional subject disciplines. As we digitise we introduce OCR technology and also hyperlinking. As newer

generations of search engines emerge, what newer inter-relatedness of individual information objects will emerge? What opportunities may arise for a hyperlink utopia as envisioned by Ted Nelson (1987)? One where anything has the possibility of being linked to anything else through individually initiated, post-coordination of search terms, where perhaps the existence of pre-coordinated texts is nothing more than a mindset generated by older technologies and their limitations. In such a scenario might all of our current software tools and all of our current digitisation projects combined to produce a fusion and a levelling of knowledge relationships? To do so would re-engineer the print block, those millions of paper blocks that hold and effect the modelling and the accessibility of knowledge as we know it.

The Burial of the Book

Indeed, perhaps it sounds like a chapter heading from a work of fiction, but is such a thought too far from reality already? There are interesting projects and theories currently in existence which are worth considering in this regard and they are to be found in the literature of repository storage. Take Schottlaender's (2004) model for degrees of library storage and accessibility of the printed word i.e. light, dim and dark. In the case of 'light' put the library which has all of its collections on open access. In the case of 'dim' put libraries which have some of their collections off-access in depositories; though retrievable. In the case of 'dark' put those collections which are placed in long term storage, totally off-access and intended to stay in safe keeping as a back up resource in the event of a disaster in the digital knowledge world. Such collections are deep archived. As the digital world impacts more and more on the economic basis for holding large quantities of printed stock on open access; as libraries begin to place greater quantities of stock than ever before in remote storage; as demand for recalls from such stock fades out; will the issue of deep storage at national level present itself as a means of allowing financial planners to rationalise ongoing expenditure? Might we bury the book in environmentally controlled vaults because it no longer performs its socially intended function, being instead regarded as just one mode of archival preservation for a portion of the world's intellectual heritage, one captured over a period of about a thousand years from the mid 15th to say the mid 21st century.

A Digital Scenario of the Future

As our fictional scenario of the pool and the grid emerges, let us consider two stages of discovery relating to the uniqueness of knowledge/information. In the first stage all knowledge stored wittingly and unwittingly in physical media has been identified, digitised and described in metadata format. All meanings, all ontologies have been standardised and rationalised. Therefore there is universal integrity of all stored knowledge and meaning captured since the beginning of the process of human recording of knowledge/information. In the course of the digitisation process and its evolution it has been discovered that the quantity of unique, original thought and documentation within the corpus is a significantly small part of what exists and that the lion's share is in fact replication, re-verbalisations due to conventions of scholarly writing; market demand for reprints, digests, editions; a variety of distribution/dissemination formats. Many millions of gigabytes are dedicated to the storage of this replication and information managers/economists decide that it is of no value to continue holding onto it. Therefore it is siphoned from the pool and either compressed for archival preservation as a record in the history of pedagogy or else destroyed. Once this has happened, as the second stage in the emergence of the pool, knowledge/information is regarded as a purer thing, it has regained some of the mystique it held in the early years of literate society.

Power

In whose garden might the pool reside? In whose landscape might the waters which continue to feed it flow? There are academic players in this game of the future. There are corporate players. There are governments, nations, empires in its history. There have always been those who are literate and those who are not so. There has always been an information rich and an information poor. The degrees involved have always varied with changes and movements in the definitions of societies and civilisations. There have always been transfers and dilutions as new torches of civilisation blazed while old ones dimmed. There have always been dreams of utopia. Is the age of digital knowledge to become this dream, one of universal access, one of universal possession, one of universal ownership? Would it be truer to say that such a dream will always remain a fiction, that there will always be multiple, segmental ownership and access. If so how political might the issue of market share of the pool become and how might that effect the humanitarian ideals and scholarly freedoms of expression so cherished by western society through its tradition of the book? The relationship between the pool and the seeker may not be a democratic one.

The existence of an overwhelmingly vast quantity of knowledge/information does not necessarily imply that the future will have use for it. At any one time what quantity of knowledge/information is of use to a society, to a generation of people, to any one cultural phase in the growth of a society? A mid 1850s trade directory is a fascinating text for a historical geographer but for society in general it has long lost its currency. How much of the flood of exabytes quantified in the SIMS report is it necessary to preserve to facilitate the future's memory of the past? How important is the memory of the past and in how many aspects of everyday life does it impact on us from self-defining cultural awareness, to legal precedent, to audits of expenditure? Therefore will the pool contain everything we currently considering preserving and migrating to it or will its waters be filtered and if so through what criteria?

Book Collection Survival in the Digital Age

It is interesting to note the emergence in recent years of major collection mapping initiatives; usually co-operative national efforts. It is also interesting to note an increasing interest in gathering / creating special collections. To an extent libraries have always realised the importance of special collections as part of their identity, as an expression of intellectual wealth. Mapping projects provide the opportunity to gather information about these holdings and to standardise description of their content. Special collections is a very varied term when one tries to define its precise meaning. It encompasses collections built in-house to a specific theme, collections built privately and then acquired for preservation, collections which were once common working collections but which have become rare or unique due to time and natural wastage. What is 'special' about these collections is that they have become so because of the integrity which defines them. The 'integrity' is what makes them into national monuments. Because of this they are given special consideration in terms of preservation programmes and frequently they are used as teaching aids in a manner similar to the use of museum items be they artifacts, geological or zoological samples. This different focus, this role, in the use of the book as an object of text transmission is perhaps one reason why such collections may survive the burial of the book, and the digital age. But if so the landscape in which they survive and the repositories in which they are accessed may be very different from we know. If one visits old libraries today, for example surviving monastic collections, surviving 17th century libraries, 18th century gentlemen's libraries where the books remain in the original rooms, on the original shelves, with their original finding aids what are you witnessing? Collections which are obsolete, stagnated within their timeframes where one is exposed to the mental worlds of other times, sitting there wanting to speak to the next reader. Such places speak to the historical imagination as much as ruined castles and abbeys. Perhaps as we continue to build and acquire special collections for the future we are creating a facet of the

physical landscape of 2,500 one where the monuments of intellectual history are libraries of special collections; collections which have survived not as sources of knowledge but rather as the tangible relics of former times.

Conclusion

Therefore are libraries, as we know them, a vanishing entity? It seems likely that the answer is yes but in what timeframe is difficult to say and how this might be engineered is perhaps for the next two to three generation of society to decide. The reasons for vanishing may well be dictated by market forces rather than intellectual need; a consequence of the commercialisation of knowledge rather than being a scholarly choice. So to end with a few questions. In the development of our collections and in our discussions of preservation standards and policies, where should librarianship position itself in the world of the digital age? Should it split itself into custodians of the print tradition on the one hand and creators, mentors of the digital pool on the other? Should it flow with the tide of change or become the instrument of its future design? Perhaps there is much to learn about this issue through re-evaluation of those old collections I mentioned above; what they have to tell us about library collections as the medium of the message rather than the message itself.

References

This essay is an exercise in speculative writing. It is not a research paper. It is derived from ideas generated by general reading in a number of sources, some of which I have listed below. The writings of Ted Nelson as well as Vannevar Bush's seminal paper (*Atlantic Monthly*, 1945) are an underlying springboard. But to this may be added Google's announcement in December 2004 regarding the proposed digitisation of the book collections of the libraries at Stanford, Harvard, Michigan and Oxford universities; along with the New York Public Library. The result is an estimated online library of 30 million volumes. Could the meshing of such a resource with the vast quantities of existing digital knowledge - which is growing exponentially through retrospective digitisation of physical documents coupled with e-born ones, become the beginnings of our pool?

- Berners-Lee, T. (2000). *Weaving the Web: The original design of the World Wide by its inventor*. New York: Harper Business.
- Bush, V. (1945, July). What We May Think. *Atlantic Monthly*.
- Febvre, L. & Martin, H-J. (1997). *The coming of the book: The impact of printing 1450-1800*. London:Verso.
- Gill, T. (2004). Building semantic bridges between museums, libraries and archives: The CIDOC conceptual reference model. *First Monday*, 9 (5), May. Retrieved on 05/01/2005 from http://firstmonday.org/issues/issue_9_5/gill/index.html
- Heaney, M. (2000). An analytical model of collections and their catalogues. UK Office for Library and Information Networking.
- Lowenthal, D. (1985). *The past is a foreign country*. Cambridge :Cambridge University Press:
- Lyman, P. & Varian, H. R. (2003). How Much Information. Retrieved on 05/01/2005 from <http://www.sims.berkeley.edu/how-much-info-2003>
- Nelson, T. H. (1987). *Computer Lib / Dream machines*. Washington: Microsoft Press.
- Nichols, S. G. et al. (2001). *The evidence in hand: Report of the task force on the artifact in library collections*. Chicago: Council for library and Information Resources:
- Schottlaender, B. (2004). You say you want an evolution. *Library Collections, Acquisitions & Technical Services*, 28 (1), Spring, 13-24.

Summerfield, M., Mandel, C. & Kantor, P. (1999). The potential of online books in the scholarly world. Retrieved on 05/01/2005 from the Columbia University Online Books Evaluation Project. Columbia University: USA. <http://www.columbia.edu/cu/libraries/digital/texts/about.html>

Reilly, B. & DesRosiers, B. (2003). *Developing print repositories: Models for shared preservation and access*. Chicago: Council for Library and Information Resources.

Biography



J.P. McCarthy is an Assistant Librarian who has worked in the profession for the past 27 years. At the beginning of his career he worked in the area of rare books, special collections, manuscripts and archives. Subsequently he occupied a number of roles ranging from cataloguer, automation librarian, project manager off-site storage, Arts and Multimedia Librarian. He is currently involved in establishing a remote storage facility for the Boole Library.

Derived from his interest in the history and heritage of Librarianship, he has an ongoing interest in the future of libraries and in the impact new technologies may have on that future. In this context, he is also interested in how the relationship between libraries and their clientele may develop and change.