This book examines various views and perspectives on digitisation. As Simon Tanner, Director Digital Consultancy, King’s College London says in the Foreword:

“Digitisation has become a cultural, scholastic, economic and political imperative and raises many issues for our consideration.”

Furthermore, that the book:

“...seeks to address and answer some of the big questions of digitisation... It succeeds on many levels...”

There are 22 contributors in the book, all experts in their fields. The book is divided into six parts:

Part 1: ‘Background and Overview to Digitisation and Digital Libraries’
Part 2: ‘Digitisation and Higher Education’
Part 3: ‘Digitisation and Inequalities’
Part 4: ‘Digital Libraries, Reference Services and Citation Indexing’
Part 5: ‘Digitisation of Rare, Valued and Scholarly Works’
Part 6: ‘Futuristic Developments of Digitisation’

Topics covered include electronic theses, search engine technology, digitisation in Africa, citation indexing, reference services, the Scholarly Publishing and Academic Resources Coalition, new media and scholarly publishing. The final chapter explores virtual libraries, and poses some interesting questions for possible futures.

The book will be of particular interest to information professionals, educators, librarians, academics and IT and knowledge experts.

Ruth Rikowski concludes by indicating that:

“...hopefully, the book will provide a source of inspiration for further research, leading to some more effective ways to proceed with the digitisation process. Also, that it will be possible to do this within a framework that can be used for good rather than ill, and for the benefit of many.”
Digitisation Perspectives
EDUCATIONAL FUTURES
RETHINKING THEORY AND PRACTICE
Volume 46

Series Editor
Michael A. Peters
University of Illinois at Urbana-Champaign, USA

Editorial Board
Michael Apple, University of Wisconsin-Madison, USA
Miriam David, Institute of Education, London University, UK
Cushla Kapitzke, Queensland University of Technology, Australia
Simon Marginson, University of Melbourne, Australia
Mark Olssen, University of Surrey, UK
Fazal Rizvi, University of Illinois at Urbana-Champaign, USA
Linda Tuahwai Smith, University of Waikato, New Zealand
Susan Robertson, University of Bristol, UK

Scope
This series maps the emergent field of educational futures. It will commission books on the futures of education in relation to the question of globalisation and the knowledge economy. It seeks authors who can demonstrate their understanding of discourses of the knowledge and learning economies. It aspires to build a consistent approach to educational futures in terms of traditional methods, including scenario planning and foresight, as well as imaginative narratives, and it will examine examples of futures research in education, pedagogical experiments, new utopian thinking, and educational policy futures with a strong accent on actual policies and examples.
Digitisation Perspectives

Edited by

Ruth Rikowski
Chandos Publishing, Oxford
# TABLE OF CONTENTS

Acknowledgement ........................................................................................................ vii

Foreword .................................................................................................................. ix
    Simon Tanner

Introduction .............................................................................................................. xi
    Ruth Rikowski

**Part I: Background and Overview to Digitisation and Digital Libraries**

1. The Rise of Digitization: An Overview .......................................................... 3
    Melissa M. Terras

2. Digital Libraries and Digitisation: An Overview and Critique .................. 21
    Ruth Rikowski

3. Digital Knowledge Resources ..................................................................... 43
    M. Paul Pandian

4. Digitisation: Research, Sophisticated Search Engines, Evaluation – All
   that and more .............................................................................................. 65
    Ruth Rikowski

**Part II: Digitisation and Higher Education**

5. Improving Student Mental Models in a New University Information
   Setting ............................................................................................................... 89
    Alan Rosling and Kathryn Littlemore

6. Electronic Theses and Dissertations: Promoting ‘Hidden’ Research ......... 103
    Susan Copeland

7. Learning Systems in Post-Statutory Education .......................................... 115
    Paul Catherall

8. Going Digital: The Transformation of Scholarly Communication and
   Academic Libraries ....................................................................................... 131
    Isaac Hunter Dunlap

**Part III: Digitisation and Inequalities**

9. Hegemony and the Web: The Struggle for Hegemony in a Digital Age ..... 147
    Tony Ward
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Digital Library: An Opportunity for African Education</td>
<td>167</td>
</tr>
<tr>
<td></td>
<td>Dieu Hack-Polay</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Critical Perspectives on Digitising Africa</td>
<td>179</td>
</tr>
<tr>
<td></td>
<td>Leburn Rose</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Part IV: Digital Libraries, Reference Services and Citation Indexing</strong></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Digital Library and Digital Reference Service: Integration and Mutual Complementarity</td>
<td>191</td>
</tr>
<tr>
<td></td>
<td>Jia Liu</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>The New Generation of Citation Indexing in the Age of Digital Libraries</td>
<td>213</td>
</tr>
<tr>
<td></td>
<td>Mengxiong Liu and Peggy Cabrera</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Part V: Digitisation of Rare, Valued and Scholarly Works</strong></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Building the Virtual Scriptorium</td>
<td>229</td>
</tr>
<tr>
<td></td>
<td>Tatiana Nikolova-Houston and Ron Houston</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>SPARC: Creating Innovative Models and Environments for Scholarly Research and Communication</td>
<td>245</td>
</tr>
<tr>
<td></td>
<td>Heather Joseph</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Impacts of New Media on Scholarly Publishing</td>
<td>251</td>
</tr>
<tr>
<td></td>
<td>Yehuda Kalay</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Part VI: Futuristic Developments of Digitisation</strong></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Meeting and Serving Users in their New Work (and Play) Spaces</td>
<td>267</td>
</tr>
<tr>
<td></td>
<td>Tom Peters</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Virtual Libraries and Education in Virtual Worlds: Twenty-First Century Library Services</td>
<td>275</td>
</tr>
<tr>
<td></td>
<td>Lori Bell, Mary-Carol Lindbloom, Tom Peters and Kitty Pope</td>
<td></td>
</tr>
<tr>
<td>Conclusion</td>
<td></td>
<td>287</td>
</tr>
<tr>
<td></td>
<td>Ruth Rikowski</td>
<td></td>
</tr>
<tr>
<td>Contributors</td>
<td></td>
<td>289</td>
</tr>
<tr>
<td>Index</td>
<td></td>
<td>295</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENT

I would like to thank my husband Glenn John Rikowski for his help in the final stages of the preparation of the manuscript for this book, and for his wonderful support in general.
Digitisation Perspectives asks and answers some deeply fundamental questions about digitisation. There are 22 contributors and this work takes a wide ranging tour across the sector. It starts by considering a ‘Background and Overview to Digitisation and Digital Libraries’. This is followed by ‘Digitisation and Higher Education’ and considerations of ‘Digitisation and Inequalities’. Latter sections cover ‘Digital Libraries, Reference Services and Citation Indexing’; the ‘Digitisation of Rare, Valued and Scholarly Works’ and the ‘Futuristic Developments of Digitisation’.

The fundamental questions asked and answered in this work are:

– How do we establish and create digital resources?
– Once created, how do we establish access to those resources online?
– How do we maintain those resources over the long term?
– For whom are we doing all this digital activity, and
– How do we address inequalities?

In overview these seem simple enough questions, but at any level of granularity they become complex, interwoven and deeply faceted.

To illustrate this, consider a quick assay into the world of rare manuscript digitisation. This seems straightforward in its surface justification but becomes more complex as one delves deeper. Rare manuscripts, especially Anglo-Saxon manuscripts, are often works of great intrinsic, intellectual and material value that are a source of fascination to many scholars and often the lay-person as well. And yet, they are fragile, rare and thus should not be handled too often, exposed to rough treatment or large variations in temperature and humidity. So digitisation seems an obvious and sensible solution to providing access to these rare and valued materials. We will digitise them, get the resulting images on to the Web and scholars and the lay-person can have access. In this way we can disseminate their wonder whilst preserving the original. What could be simpler or more straightforward?

As we delve deeper we see that digitisation is not such a neutral act. The process of digitisation, even when done carefully may itself become the harshest treatment the manuscript has ever undergone. Sometimes these items of great rarity have never had every page turned and yet with digitisation we will affect such a handling. Once imaged, there are many questions as to how each image should be labelled, described and indexed. Even something as prosaic as filenaming can become the source of debate and disputation in the effort to seek the best digital preservation policy. This does not even begin to address the complexity of gaining an online Web accessible version of the manuscript and addressing the user centred questions of what mode that access should take. Should it be a fancy page turning facsimile interface, or just plain images in a linear order? And that order in which they are arranged, should it be that established by a venerated scholar, such as M. R. James, or could the order be dictated by the modern physical arrangement of the manuscript pages as digitised? Or should both views be facilitated?
FOREWORD

And are digitised versions of rare manuscripts just for post-graduate scholarly use? Maybe, if they were at least calendared then a wider educational purpose might be served? But who will do this and what perspective will that interpretation take? What tools will be maintained and developed for educational use, what virtual learning environments or virtual manuscript study tools exist or could be brought into existence to serve scholarship and education? Further, will the lay-public have access and will that access be supported by guides and translations and tools to help them to engage with the content in meaningful and fulfilling ways? And will access be free or a charge levied?

It is also worth considering whether such a Westernised view of history as might be made available through a digitised Anglo-Saxon manuscript is the most appropriate expenditure of time and resources in a world of information. Would expending effort here reduce or increase the digital divide? Does it help to address the real and burgeoning needs of the developing world to spend more resources on the digitisation of a single manuscript collection than has been spent on digitising the papers and life of Archbishop Desmond Tutu in South Africa? Similarly, more has been spent on manuscripts than on distributing digitised health information on water purity and safety in Bangladesh to reduce the death rate from diseases such as Typhus. There is no neutrality in digitisation, it is beset with ethical choices and moral issues. For instance, should we charge for access to the online resources? Are we making it impossible for some countries to engage in the scholarly debate on an equal footing with their peers if we do not embrace Open Access? And yet, if we do not charge, can the online resource remain sustainable in the long term?

So, no act in this digital domain is neutral and without consequences that as Ted Nelson said “intertwingle”. It is a jungle out there, with many ecosystems competing for dominance. At its heart are human desires and these are defining the value we place upon information and the needs for digitisation to be achieved in our time. Digitisation has become a cultural, scholastic, economic and political imperative and raises many important issues for our consideration.

Thus, I welcome this book, Digitisation Perspectives, edited with great verve by Ruth Rikowski. It seeks to address and answer some of the big questions of digitisation. It succeeds on many levels, not least because it respects the deeply interlocking nature of the subject area, with high quality insights from so many experts covering so much ground.

Simon Tanner
Director, Digital Consultancy in the Centre for Computing in the Humanities
King’s College London
INTRODUCTION

I am pleased and proud to introduce my second edited collection and my third book. The subjects of all three books (globalisation, knowledge management and digitisation) overlap and interweave in many ways. In my single-authored book on globalisation, published with Chandos (2005) I explored and developed my political and theoretical positions and passions and related this to my practical library and information experience. Coupled with my absorbing interest and fascination with complex topics I then placed all this within a global/international framework. In my second book, an edited collection on knowledge management (KM) (also published with Chandos, 2007), I changed tactic. This book housed a collection of different voices and perspectives on KM, all experts in their fields, on a diversity of topics including knowledge management and wisdom; intangible value; leadership in the knowledge revolution; the role of the library in knowledge management; accessibility issues for web-based information systems; culture and knowledge management and thermodynamics and knowledge. I concluded the book with my own clear theoretical and political perspective and position.

This book on digitisation provides yet another slant; it is a collection, once again, of different expert voices and perspectives, on a wide variety of topics, this time within the broad subject matter of digitisation. However, the political dimension is kept lower key in this volume. In fact, in my second chapter in the book (Chapter 4) I end with emphasising the wonder, importance and beauty of art, which if approached in a certain way can, I think, sometimes take us beyond the confines of the political scene (albeit we all still have to work and operate within the political system, of course).

All three of my non-fiction books cover topics that are very important today, both within the library and information profession, within education and also within the wider community in general. We now need to think and operate within a global framework in many ways. Sharing, exchanging and managing knowledge (knowledge management) is very important in this global framework, as of course, is the digitisation of this material.

Digitisation Perspectives covers a wide range of important topics within the broad topic of ‘Digitisation’. It is divided into six parts. Part 1 provides a ‘Background and Overview to Digitisation and Digital Libraries’. Part 2 focuses on ‘Digitisation and Higher Education whilst Part 3 is on ‘Digitisation and Inequalities’. Part 4 covers ‘Digital Libraries, Reference Services and Citation Indexing’; Part 5 examines ‘Digitisation of Rare, Valued and Scholarly Works’ and Part 6, the final section, explores some of the ‘Futuristic Developments of Digitisation’. There are 22 contributors altogether in the book; all experts in their fields. Quite a number of books have now been published on this important topic, but hopefully, this book brings something of a different, interesting and refreshing outlook to the subject matter. The book builds on the special issue on ‘Digital Libraries’ that I co-edited with Isaac Hunter Dunlap for the refereed international ejournal,
INTRODUCTION

*Policy Futures in Education* in the first quarter of 2008 (see [http://www.wwwwords.co.uk/pfie/content/pdfs/6/issue6_1.asp](http://www.wwwwords.co.uk/pfie/content/pdfs/6/issue6_1.asp)). The book is being published with Sense Publishers at the initial invitation and suggestion of Professor Michael Peters at the University of Illinois. Professor, Educational Policy Studies at Urbana-Champaign.

There are four chapters in the ‘Background and Overview’ section. The first is by Melissa M. Terras, and is entitled ‘The Rise of Digitization: an overview’. Terras considers digitisation within three periods; the early years before the World Wide Web, the 1990s and post 2000. As Terras says, in the 1980’s:

…interest in the application of digitization and image processing in the arts, humanities, library and archive sector began to grow.

There was a significant growth in digitisation in the 1990s, which was due to a variety of forces, such as the increase in the performance and availability of new networked technologies and in the awareness of the possibilities that the new technologies could bring forth. Whilst in the post 2000 period we have witnessed projects such as the JISC (UK Joint Information Systems Committee) digitization programme, and digitisation work being undertaken by commercial firms, such as Google and Microsoft. Terras concludes by saying that:

There are interesting times ahead for the information professional in dealing with the wealth of digital information now available, keeping up with the changing nature of Internet technologies, and coping with resulting changes in user expectations and needs, to create digital libraries which can exploit the potential of this groundswell of digital – digitised – content.

This is followed by an overview and critique of digital libraries and digitisation which I have written. I consider topics such as the advantages and costs of digitisation; traditional, digital, hybrid and virtual libraries; the library community and digitisation; ebooks; digital libraries on a global basis; gender issues, IT and digitisation and the social and political implications of digitisation. Some of the advantages of digitisation outlined include the ability to be able to search, browse and compare a variety of material; developing digital surrogates of rare or fragile original objects; bringing collections together in a virtual, digitised form, that might not otherwise be bought together and raising the profile and prestige of an organisation. Various digital projects are considered, such as the Oxford Digital Library project and Alouette-Canada, a Canadian national digital online project. In regard to the social and political implications, as I say this:

…includes issues around the digital divide; the exploitation of workers in the developing world; the possible loss of a sense of community; the concept of ‘self’ in the IT age; …and potential health risks through over-use of e-reading (backache, eye strain etc.). …issues related to the likelihood of less face-to-face teaching … (due to the increase in e-learning, easy access to digitised information, etc) and those related to the ever-changing nature of technology…

xii
I conclude by saying that:

Digitisation and digital libraries offer many new and exciting ways for us to be able to access material, but we also need to exercise some caution, and not just jump on every digital bandwagon.

The third chapter is by M. Paul Pandian and is entitled ‘Digital Knowledge Resources’. Pandian provides an historical digital tour, whilst also demonstrating the fast pace of the digital transformations that have taken place over the last few years. Pandian considers emerging technologies, and how these technologies have changed how librarians, information professionals and researchers organise, search, retrieve and preserve information. Pandian also compares and contrasts the digital library with the traditional library, and the valuable role that the traditional library still plays today. In regard to the global nature of digitisation, Pandian says that:

Libraries of all types and in all settings are developing a global vision of international networked collections and services. This model views libraries as both providers of worldwide knowledge resources for gateways for users to knowledge which is increasingly electronic in form. This transformation requires recognition of important revolutionary changes, which are transforming collections, information services and the working relationships among libraries.

The last chapter in the first section is another one by me, and is entitled ‘Digitisation: research, sophisticated search engines, evaluation – all that and more’. As the title implies, this chapter examines a number of different areas of digitisation, in more depth. It explores topics such as digitisation research and development; electronic theses; ‘googlisation’ and search engine technology. Also explored are: data mining large digital collections and application programming interfaces; the development of digital libraries for the scientific communities in China; the value of digital preservation; the evaluation of digitisation projects and Arms ‘viewpoint analysis’ of the digital library.

In regard to sophisticated search engine technology, I focus on the work of Summann and Lossau; Brophy and Bawden; Xie (who compared online database systems, Dialog and Factiva with three different types of web search tools – a search engine, a directory and a meta-search engine), and the ‘information model for digital libraries’ as outlined by Lagoze, Krafft, Payette and Jesuroga. Whilst application programming interfaces assist with data mining of large digital collections, and Cohen’s Syllabus Finder, in particular, is considered. I conclude with Arms ‘viewpoint analysis’, which looks at digital libraries from three different viewpoints – the organisational view, the technical view and the view of the user. Arms argues that greater consideration should be given to the users viewpoint. Whilst this might seem rather romantic, I make the point that on one level such an approach can, perhaps, take us beyond the rather limited political and economic scene that we can sometimes find ourselves embroiled in today.

Part 2 examines ‘Digitisation and Higher Education’. The first chapter in this section (Chapter 5) is on ‘Improving student mental models in a new university information setting’ and is by Alan Rosling and Kathryn Littlemore. Their chapter
INTRODUCTION

outlines the culture and ethos of the university and the students that they encounter and interact with there within the broad context of digitisation. They ask:

How well are our students able to make use of the digital resources we provide, and how, as teachers, do we in fact teach information literacy skills in the digital setting?

They comment on the ‘Google Generation’ and ‘Digital Natives’ and ask:

The label ‘Google Generation’ implies that these individuals have a natural affinity with anything online or digital, that they are completely comfortable with existing and new technologies, and hence, their mental models of digital resources are well developed – and easily translated into digital libraries. Is this a fair and realistic assessment of students?

The chapter focuses in particular on ‘mental models’ and mental model theories. Craik (1943) was one of the first people to describe these mental models, saying that they are ‘small scale models’ of reality. Whilst Norman (1983) distinguishes between mental and conceptual models, where the mental model is what the user has in her or his mind and the conceptual model is the model that is given to the user by an outsider, such as by a designer. Rosling and Littlemore reflect on mental models saying:

So our user comes to their ‘user education/information skills/information literacy’ class bringing with them their past experiences and prior knowledge of computers and online searching and software. As they sit at the terminal they also bring with them their, often very incomplete, mental models of the information landscape; the Internet, how computers work and how they might search successfully.

Mental models and information literacy teaching is also explored, as well as the role of emotional intelligence within this setting. Rosling and Littlemore conclude by arguing that:

The digital landscape requires more from the librarian…than it did in the days of the print landscape. Teaching ‘search strategy’ skills is no longer adequate for the complex environment of the Internet and web-based search tools. Mental model research can help librarians to use more focused teaching strategies like encouraging students to explore (or even play), to target more sophisticated feedback towards specific digital problems (why searches fail, access rights, comparing systems)…If we continue to use current mental model research to make even more sense of the student searching experience then we might establish even clearer learning strategies that will move learners further up the information literacy skills ladder.

Chapter 6 explores ‘Electronic Theses and Dissertations: promoting ‘hidden’ research’. In this chapter Susan Copeland examines the development of e-theses internationally and some of the cooperative ventures that have been undertaken to
INTRODUCTION

facilitate scholarly communication and resource and knowledge sharing. This includes the development of the Networked Digital Library of Theses and Dissertations (NDLTD). As Copeland says, her chapter:

…highlights key issues associated with the creation, management and use of electronic theses and dissertations and provides information about organisations that are actively working to promote this useful source of research data. Information is provided on topics such as training, technical and administrative requirements, and the issues to address whether adopting a local, institutional, approach or participating in a national level service.

Copeland concludes her chapter on an optimistic note, highlighting the fact that although there are “Differences in terminology, procedures and requirements associated with how theses and dissertations are presented in different countries…” improvements are continually being made in this area. The NDLTD has played a particularly valuable role in this regard.

In Chapter 7 Paul Catherall provides an overview of ‘Learning Systems in Post-Statutory Education’. This includes a consideration of the origins and forms of learning systems; technical, educational concepts and approaches used, such as distributed and collaborative learning; the VLE (Virtual Learning Environment) and trends in e-learning including future technologies such as m-learning (mobile learning). Some of the challenges in the delivery of systems and software to facilitate learning in a digital environment are also highlighted. In regard to e-learning systems, Catherall says that:

Web-based learning systems provide a range of interactive functions, including communication channels, content-publishing and assessment tools. It can be seen that these systems have begun to permeate and impact on many aspects of post-statutory education, with recent web technologies allowing for closer systems integration across institutional systems, including library catalogue, student records and finance systems. In the United Kingdom, this institute-wide computing infrastructure is often considered a holistic learning environment.

Meanwhile, Isaac Hunter Dunlap delves into the realms of ‘Going Digital: the transformation of scholarly communication and academic libraries’. As Dunlap says:

This chapter provides both historical perspective and a forward-looking examination into how academic libraries are transforming themselves to both cope with, and help shape, unprecedented transitions in scholarly research and communication.

He notes that, in general, digital libraries:

…have become a complex, interactive and interdependent network of resources, user services, social technology interactions, databases, virtual meetings, gateways, ‘real time’ online instruction opportunities and content management systems.
INTRODUCTION

University libraries are now embracing the emerging technologies and Dunlap considers how structures and functions within the university library are being repositioned and transformed so that they can fit more effectively into this digital age.

Within this, he explores a number of different areas including the hybrid library; financial constraints; the Open Access movement; the acquisition of digital collections; social networking technologies and the digitisation of special collections.

There are three chapters in Part 3 ‘Digitisation and Inequalities’. The first of these (Chapter 9) is entitled ‘Hegemony and the Web: the struggle for hegemony in a digital age’ and is by Tony Ward. Ward draws on his practical experience as a teacher within a critical pedagogical approach. He considers topics such as the digital divide; the promise of the Internet (the fact, for example, that as Ward says “The Internet is seen as a space where everybody can have a voice, can be heard and can express their own realities…”); scarcity, and the creation and control of knowledge and information systems within this; education as a commodity and the modern media. Ward concludes by saying that it is clear to him that the Digital Divide:

…will never be “bridged” or “closed” as long as we have an economic system that operates on the basis of exclusions...The modern digital media holds an enormous sway over the public mind and in the hands of powerful groups and individuals poses a formidable obstacle to the emergence of a popular democracy. But there have also been many gains in that time. The struggle for hegemony will be ongoing and endless.

And his final thought is this:

The question remains whether...shifts in the use of technology will ever provide the access to public voice that the truly oppressed can use for their own emancipation. The movement for open access may be one of the most critical in our age.

Ward’s desire for a fairer and a kinder world is self-evident.

This is followed by two chapters that explore digitisation in Africa specifically. Chapter 10 is by Dieu Hack-Polay and is entitled ‘Digital libraries: an opportunity for African education’ and Chapter 11 is by Leburn Rose and is entitled ‘Critical Perspectives on Digitising Africa’. Hack-Polay, considers the many benefits that the development of digital libraries can bring to Africa. These include the fact that digital libraries can “…provide a forum for inter-institution communication and updating on the currency in different subject areas”; that data can be retrieved more easily; and that it provides great opportunities for the dissemination of both unpublished and published material in Africa. Hack-Polay explores some of the digital initiatives that are taking place in Africa, such as the African Online Digital Library (AODL). Furthermore, the variety of sub-cultures and different languages and traditions is emphasised and the helpful role that digitisation can play here. Hack-Polay says:

…digital libraries could be a platform where the education system makes available notions about country or culture-specific histories, stories, good
practice models in terms of educational, political, scientific and technological development.

Some of the constraints for the effective development of digitisation are also outlined; such as cost, logistics and manpower constraints. Never-the-less Hack-Polay concludes on an optimistic note, saying that:

Digital libraries could ultimately, more than any other means before, aid literacy development in Africa and in higher education, foster more quality academic output and research.

Meanwhile, Leburn Rose in Chapter 11 is somewhat more sceptical and concerned about how western approaches to digitisation can and do impact on developing countries in Africa. The infra-structure in Africa is very different to that in the developed countries. He refers to the ‘dual nature’ of technology (including digitisation itself), arguing that it is both ‘instrumental’ and ‘volitional’ at the same time. But as he says, the ‘volitional perspective’ is western-based.

The volitional perspective [the will to power] on technology is quintessential Western, and presents a distorted view of technology as something that is being created for and imposed upon.

He says that a framework for ‘authentic digitisation’ needs to be developed in Africa arguing, for example, that:

Those that seek to sponsor authentic digitisation should focus on strategies that build upon and engender social well-being and are contextually sensitive to the relevant domains of practice.

He then interestingly applies the ‘free-body diagram’ (a concept adapted from his engineering background) to the picture, whereby a part of the technology is isolated so that it can be ‘subjected to analysis’. As he says:

The conclusion here it not surprising: that the problems which technology attempts to resolve and the outcomes sought are contingent upon the surroundings of which the technology may be a part. One cannot simply export a technology into an unfamiliar context and assume that it will yield similar results to those achieved in the original setting.

Thus, in essence, western digitisation programmes cannot simply be imposed on Africa, or indeed, on the developing world in general; rather, they need to be adapted to the social and economic conditions within these countries, and to their particular infrastructures.

Part 4, ‘Digital Libraries, Reference Services and Citation Indexing’ returns to some more traditional library services and approaches, and how these are now being developed and changed through and with digitisation practices. Through this we can aim to retain the quality and worth of the traditional aspects of librarianship and the skills that have been learnt and developed, and then build and develop all this within a digital environment. The first chapter in this section is about digitising
INTRODUCTION

the reference service. The chapter is by Jia Liu, and is entitled ‘Digital Library and Digital Reference Service: integration and mutual complementarity’. Through a series of case studies, Liu argues for a “three-dimensional integration”, operating around the digital reference service, the digital library, and areas of overlap and integration between the two services. She notes the fact that whilst the term the ‘digital library’ is generally known and widely used, the ‘digital reference service’ is also known under various other terms, such as the ‘virtual reference service’; ‘online reference service’; ‘electronic reference service’ and ‘ask a/the librarian service’. Liu argues that:

It is logical, reasonable and necessary for the digital reference service to take the digital library as part of its reference collection.

Some of the work that has been undertaken on the creation and development of digital reference services is considered. The University of California library, for example, has been a pioneer for providing a remote reference service, and in 1997 the Science Library librarians there implemented a pilot ‘telereference’ project. This project focused on topics such as desktop videoconferencing for conducting reference interviews with student users at a remote location. Liu concludes by saying that:

Both the digital library and the digital reference service are products under the networked environment and evolve with the developments of information and communication technology (ICT)...Without any doubt, the digital library and the digital reference service have a close relationship and need each other...

The other chapter (Chapter 13) in this section is a detailed analysis of citation indexing; it is entitled ‘The New Generation of Citation Indexing in the Age of Digital Libraries’ and is by Mengxiong Liu and Peggy Cabrera. Liu and Cabrera review developments in web-based citation indexing and conduct a case study into three major citation search tools; these being Web of Science, Scopus and Google Scholar. They found that:

...none of the three tools can satisfy all of a researcher’s citation tracking records. Web of Science showed strength in providing citing references to traditional academic journals while Scopus performed better in providing citing literature for more current articles. Google Scholar returned a significant number of non-traditional citing references. With its advantage of free availability via the Internet, Google Scholar is an important compliment to WoS and Scopus.

Following on from the various evaluations and comparisons of citing reference services, Liu and Cabrera conclude that

...there is no single solution for a complete citing reference search without a Universal Citation Digital Library.

Part 5 explores ‘Digitisation of Rare, Valued and Scholarly Works’ and there are three chapters in this section. Tatiana Niklova-Houston and Ron Houston have
written a fascinating and absorbing chapter about the digitisation of ancient manuscripts, which is entitled ‘Building the Virtual Scriptorium’. Whilst digital preservation, in general, is seen to be quite worthwhile as an aim, the work on the digitising of ancient manuscripts proves to be particularly challenging. Niklova-Houston and Houston address some of the major challenges and obstacles for the digitisation of ancient manuscripts. They argue that a clear digitisation programme needs to be adopted, which is very mindful of emerging technologies and evolving industry standards. They say that:

Manuscripts, archives, and early printed books contain a documentary record of the foundations of human knowledge...On the assumption that the widespread availability of knowledge benefits the human condition more than the restriction of knowledge, elements restrictive to the dissemination of manuscripts, archives, and early printed books should be overcome, and the intellectual content of such items should be available to as wide an audience as possible through the digital library equivalent of the medieval scriptorium, termed here the “virtual scriptorium.”

Their chapter includes some lovely photographs of various ancient manuscript material that have been digitised and Niklova-Houston and Houston look at three SlavMan (Slavic Manuscript) websites. One of the obstacles to digitising this material comes from the curators of traditional manuscript collections themselves. Curators ask questions such as: “How will digitizing further our secondary mission: to support the scholar?” Another obstacle, of course, is the financial one. Niklova-Houston and Houston conclude by saying that:

Knowledge is out there, stored in the pages of manuscripts locked away in dusty archives. It will remain there, unused, unless you take your digital camera, digitize the images, and post them in a web-based virtual scriptorium. You may not create the perfect library; you may not have sufficient permissions; your work soon may require formatting. Yet, whether or not you digitize, these problems will persist. Digitize now, and they may be overcome.

Heather Joseph, Executive Director of the Scholarly Publishing and Academic Resources Coalition (SPARC) outlines the work of SPARC in a chapter entitled ‘SPARC: creating innovative models and environments for scholarly research and communication’. She explores some of the main activities of this influential organisation, which takes a leading role in shaping the future of scholarly works and communication. SPARC was founded by the Association of Research Libraries in 1997, and it enhances the access and use of scholarly information and is developing alternate scholarly communication models (such as open access models) that utilises digital technology. Joseph explores three of SPARC’s major programme areas; namely, education, incubation and advocacy.

Whilst in Chapter 16, Yehuda E. Kalay, explores the ‘Impacts of New Media on Scholarly Publishing’. He considers how these ‘new media’ (such as blogs, wikis, open source, podcasts and the World Wide Web) have dramatically altered academic scholarship on many different levels. The focus is on scientific publishing and the
chapter summarises some of the key results of a workshop that was held in the University of California Berkeley in June 2006 that was organised by the Center for New Media (CNM) and supported by Elsevier, the leading publisher of scholarly journals in the field. This was the first of a number of planned workshops. The workshop asked questions such as: “How will scientific publishing be affected by New Media?” Kalay suggests that the New Media technology transformation should be viewed from a “horseless carriage” point of view:

…not merely as a technological revolution, but as a wholesale social reorganization, affecting the very core of human knowledge, the modes of its production, validation, and dissemination, and shifting the locus of knowledge production from institutions to individuals, complete with relocation of authority over knowledge authentication and valuation.

Part 6 moves on to look at ‘Futuristic Developments of Digitisation’. Tom Peters contribution here is in a chapter entitled ‘Meeting and Serving Users in their New Work (and Play) Spaces’. Peters investigates how the upsurge of personal computing devices and related growing digital information environments are creating new and varied needs and expectations. Peters argues that experimentation, inventiveness and creativity will become increasingly important within library and information services in the future, as people’s needs and demands become more and more diverse and complex. Within this remit, Peters examines the public services component of both digital and virtual libraries. He concludes by saying that in the future people:

…will be using a wide variety of personal, portable information/communication/entertainment devices. We may be facing a future in which no single portable device or even device type comes to dominate the marketplace. Libraries will be expected to deliver services to them all. [we will be sharing]…information and insights across these multiple worlds. Learning and information seeking may become significantly more participatory.

In Chapter 18 by Lori Bell, Mary-Carol Lindbloom, Tom Peters and Kitty Pope, Bell et al look at ‘Virtual Libraries and Education in Virtual Worlds: twenty-first century library services’. This highly futuristic chapter explores the virtual library through the for-profit Second Life electronic community. Second Life is a completely virtual world in cyberspace, and includes a variety of facilities, such as shops, theatres and libraries. People buy space in this virtual world and then interact. Bell et al say that:

As the use of the Internet and time spent on the Internet by individuals grows, and the use of virtual worlds like Active Worlds and Second Life increases, the library needs to have an interactive place and role in these worlds as well as a bricks and mortar space.

The chapter focuses, in particular, on a collaborative library project in Second Life that was set up by two of the authors (Bell and Pope). As Bell et al say, new participants to Second Life “…discover a vast digital continent, teeming with people,
entertainment, and opportunities.” There are over 50 libraries in Second Life. The Cullom-Davis Library at Bradley University in Peoria, Illinois has set up a replica of the real-life library, in Second Life, for example. Bell et al note that challenges for libraries in Second Life include staffing, funding and sustainability. They conclude by saying that:

We believe that virtual worlds such as Second Life are the next phase of the Internet...[libraries] need to create an effective presence in virtual worlds to investigate what kinds of library services will work and how they can work together to benefit and serve users. It is also important that libraries work closely with other educators and collaborate with groups. No one library can do everything themselves – collaboration will be key in establishing virtual world presences for libraries.

It will be interesting to see how virtual libraries progress and shape-up in the future, and what sort of impact they are likely to have/will have on the wider society.

Hopefully, this varied collection of works on the broad topic of ‘Digitisation’ will encourage thought and reflection as well as being informative, enthralling and engaging. Also, that we can aim to focus on the positives of digitisation and endeavour to make digitisation work for the good and the benefit of the majority, rather than just for a select few.

REFERENCES


Ruth Rikowski
Chandos Publishing, Oxford
May 2010
PART I: BACKGROUND AND OVERVIEW TO DIGITISATION AND DIGITAL LIBRARIES
Prior to digitisation, the retrieving and re-filing of the prints consumed staff time. Batches of prints to be re-filed hung round on trolleys like patients in a busy A&E Department. The sight of a pink print request form approaching the Local Studies counter caused the heart to sink - it foretold at least half an hour wrestling with ancient filing cabinets, not to mention the wait for the notorious Central Library lift. And what could be more dispiriting, for staff and customer, to discover that the prints selected from the card index were not what were wanted at all? The idea of digitising the Print Collection was very attractive. It would reduce the wear and tear on the original prints and transform public access by allowing customers to browse the whole collection (Moorhouse 2004, 62).

Digital libraries depend on digital content. There are information objects which are “born digital”, having been created by computational technologies and never existing in analogue format, but the majority of holdings offered by digital libraries were created through the process of digitization, “the conversion of an analog signal or code into a digital signal or code” (Lee 2002, 3).

The information, culture, and heritage sectors were quick to embrace digitization technologies as they became available, primarily to facilitate access to items in collections by providing them in electronic format. Producing digital versions of holdings through digitization has become an industry in itself, and there has been much effort devoted to producing guides to best practice for undertaking the digitisation of library and archive material.

This chapter provides an overview of the rise of digitization, demonstrating the phenomenal recent uptake of digital technologies within the library sector, laying the foundations for the digital libraries of the future.

EARLY YEARS: BEFORE THE WWW

Digitization may seem a recent phenomenon in memory institutes, but the current state of affairs where libraries, archives, museums, galleries, and even private collections are expected to make available their holdings in digital form follows a period of experimentation with and appropriation of available digital technologies which dates back almost forty years. Institutions began to utilize computer systems in the 1970s.
… with the creation of electronic catalogues. The analogue card catalogues that provided access to the collection items were replaced by databases. The bibliographic information on the catalogue cards was converted into database fields. The electronic catalogue provided automatic access to analogue objects such as books, works or art and documents. In the 1980s the conversion of printed source material, such as books and articles, into digital files started to become widespread. Two types of digital files were created. The first file type represents the storage medium of the original analogue document; this is often called a digital image. The second file type represents the coded content (for instance, characters and figures) of the original document, and, depending on the type of original, can be an electronic text, table, or database (van Horik 2005, 12).

In the 1980s, interest in the application of digitization and image processing in the arts, humanities, library and archive sector began to grow. Small scale, in-house projects of limited scope and interest were undertaken by individual institutions experimenting with the application of the newly available (but still expensive) technologies (González 1992), but towards the end of the decade large scale projects were launched by various institutions, including pilot projects aiming to investigate the appropriation of digital technologies to the handling of large volumes of information.

One of the first digitization projects was launched in 1984 by The National Archives and Records Administration (NARA) in Washington. Their project “Optical Digital Image Storage System (ODISS)” aimed to test the utility of digital image and optical disk technologies for the reproduction, storage and retrieval of archival documents. Digitizing 220,000 pages of documents, the ODISS project found digital representations of their holdings to be easy to use, with significant time saving for searching and retrieving information, improvement of legibility for documents digitised in high resolution, reduction of storage space, reduced handling of original documents; although they were perhaps optimistic with their predictions of the longevity of the digitised material – suggesting 100 years! (see González 1992, 1998, and NARA 1991 for an overview).

Another early adopter of digitization was the “Archivo General de Indias” in Spain, a 200 year old archive of extremely valuable documents illustrating the history of the Spanish empire in the Americas and the Philippines. Started in 1986, the pioneering and ambitious digitization project saw three public and private organizations join forces (The Spanish Ministry of Culture, IBM Spain, and the Ramón Areces Foundation) to develop a computerized information system and undertake large scale digitization of the archive holdings. By 1992 the archive had digitised 7 million pages, with 11 million completed by 1998, halting the deterioration of original papers caused by handling in the research rooms (González 1992 gives a full overview of the system and its architecture, and the processes used for digitization. González 1998 reflects on ten years of the project). Although not without its technical and legal problems, the digitised archive, which now has tens of millions of digital images of manuscript material, and electronic texts, is available for consultation in the archive itself (its online presence is rather minimal,
although the database can now be searched: http://www.mcu.es/archivos/MC/AGI/index.html).

In the library sector, various experiments were undertaken on printed library material (which has fundamentally different characteristics to archival material). One of the pioneers was the Library of Congress in Washington, which early in the 1980s launched a pilot project to look at both optical disk (for use in the digitization of general periodicals) and videodisc (for non-printed material such as photographs) – both storage mechanisms which are long since defunct (Price 1984, Parker 1985, Fischer and Swora 1986). The advantages which were found in undertaking the research included the saving in storage space compared to that taken up by the original material, although the process of digitization itself was found to be time consuming.

Another pilot project in the library area was the joint development by Cornell University and Xerox corporation, supported by the Commission on Preservation and Access, which attempted to develop a prototype for digitizing damaged, or “brittle”, books and providing high quality laser copies on request from users, at great speed, and fair prices. In December 1991, one thousand “brittle” volumes were digitised and digital editions prepared, ready to print. The project also investigated whether digitizing was a convenient alternative to the production of microfilm: finding that the time taken was actually similar, although the possibilities of access were higher (Council on Library Resources 1986, González 1992, 100). Around the same time, Yale University Library, one of the earliest developers of preservation microfilming practices,

...engaged in the first stages of a long-term and comprehensive effort to develop practical systems for the conversion of microfilm to digital image, and thereby to develop an understanding of the appropriate relationship of the two technologies in future preservation practice (Waters 1991, 1).

During this period, interest in imaging technology was also growing in the visual arts sector, with many art galleries and museums beginning to undertake projects using imaging technologies, which were becoming increasingly affordable. An attempt at cataloguing the projects that took place in the 1980s and 1990s was undertaken at the European Visual Arts Centre at Ipswich. In 1990, a meeting was hosted to discuss possible European collaboration in the form of a European Visual Arts Information Network (EVAIN) whose aim was to “maximise the potential development and effective use of image-based interactive multimedia applications within museums and galleries” (Pring 1991, iii). This led to the setting up of a database of projects in order to circulate developments and applications in the visual arts area, and in 1991, EVIAN began to publish the bi-annual database of “Image Technology in European Museums and art galleries: ITEM” (Pring 1991). Concentrating on European projects, ITEM organised a joint exchange of information with the Clearinghouse of Art Documentation and Computerization, based at the Thomas J. Watson Library of the Metropolitan Museum of Art, New York, which had documented imaging projects in art galleries and museums in the USA since the late 1980s (Stam 1987, Barnett 1991, iv). The goal of the Clearinghouse was to
“provide a gateway of information concerning relevant automated projects and activities” based on the premise that “a substantial amount of information on computerization can be shared by researchers and systems alike.” Before the days of the Internet, “this scattered but crucial information was unavailable in any one place, and required time-consuming, often redundant investigation.” (Barnett 1991, iv).

The first ITEM database detailed 112 projects, covering a broad range from inter-institutional projects funded at government level, to individual institutions undertaking digitization of their own holdings. For example, an entry on the Europe Museums Network, details its aims to

…demonstrate how new media can be developed, gain experience of multimedia design, gain experience of working with multimedia databases and computer system, and gain ideas for new services using broadband telecommunications network (Pring 1991, 6).

and to “meet the needs of museums and the telecoms industry” in “working to create new paths to art and culture for their visitors, to show that new media cannot replace art, but can lead people to it” – although there is little mention of their actual findings or recommendations. Early individual institutional digitization programmes are also catalogued, such as the Musée d’Orsay’s “Galérie des Dates” videodisc project, which produced an interactive video disc programme covering nine events in each year of sixty-eight year period. Completed in 1986, this two videodisc collection contained 15,000 slides and 14 hours of sound, allowing visitors to the gallery to explore the historical events through video clips, stills and textual information (Pring 1991, 41).

Towards the close of the 1980s,

Federal agencies …[became] increasingly interested in using digital information technologies to store large amounts of information economically and efficiently. This [was] particularly true of programs designed to provide Federal information to citizens, since a corresponding reduction in the creation of paper records could potentially reduce costs and improve the delivery of services to the public (Peterson 1994).

The National Archives and Records Administration (NARA) of the United States of America, in conjunction with the National Association of Government Archives and Records Administrators, conducted studies into digital imaging and optical media storage technologies at both State and local government levels, recommending image capture, storage, and sustainability models (NARA 1991). By the close of the decade, then, all forms of memory institutions were embracing digital technologies. In particular, imaging solutions were seen as means to make their holdings more accessible, searchable, and available, for both government bodies and the general public.

Even at this early stage, prior to the invention of the World Wide Web, what is surprising about early digitization projects was the volume of material which was digitised. Most entries in the ITEM catalogue contain thousands of digitised images: The Italian government funded “Galleria Spada” project produced more
that 5,000 images of the Spada Gallery in Rome by 1986 (Pring 1991, 85). The Arts du Costume du XXé Siècle project, at the Union Française des Arts du Costume, produced a videodisc of 33,000 images of their fashion collections, one containing collections from the 18th and 19th centuries, the other haute couture collections of the 20th Century, in 1987. La Storie dell’Arte Italiana (Pring 1991, 87) had digitised over 45,000 images of art, plus 30 minutes of video, in a project completed in 1988. The Henry Moore Sculpture Trust (UK) produced a “Sculpture Interactive” MediaBase Resource on videodisc containing 5,000 images and 70 minutes of film and radio extracts in 1990 (Pring 1991, 75). A considerable investment in time and effort was necessary to digitise such large collections, even without a ready vehicle for distribution (most digital records were consulted in the institution on stand alone machines, or towards the late 1980s, distributed on floppy disc). There clearly was a passion for, and a belief in, this new technology.

Another interesting point to note regarding many of these early projects, across all of the library, archive, and heritage sectors, is the consideration, at an early stage, of the need to instil “best practice” in the digitization process (or to determine what best practice could actually mean when applied to digitization), and the ramifications for making large amounts of material available to the general public. Many early projects were undertaken at world-leading institutions: being able to attract, or provide, the funding required for such large scale digitization. The focus on the majority of early projects tends to be large scale, with large volumes of material being captured, in the hope that Optical Character Recognition technologies would then turn the resulting images into electronic text. Much of this research was optimistic, but the trial and error approach adopted by pioneering projects, and published as bona-fide research, helped to establish many useful guidelines for subsequent digitization attempts.

However, many institutions were optimistic about the type of activity digitization could allow, and how easy it would be for the user to access these digital surrogates (without home computers, or a networked environment). Much of the rhetoric surrounding early digitization programmes was filled with

…hollow pronouncements and promises… when early experimentation with desktop technologies and remotely accessible materials for instruction and research gave senior administrators in libraries and universities, as well as funding agencies and government departments, ideas that new technology would save millions of hours of teaching time and increase academic productivity, based on the assumption that a CD-ROM of a term’s coursework could replace instructors and face to face classes. Such claims raised expectations unreasonably, and many enthusiastic “early adopters” of digital technologies discovered at great expense that there are hidden costs and pitfalls to developing and using digital content (Hughes 2004, 7).

It took a while for institutions to realise that digitization is a costly and time consuming exercise, which will not reap financial rewards but provides benefits for users, expanding skills, expertise and services, whilst requiring ongoing maintenance, development, and funding.
It is also interesting to note that many early digitization projects were funded by, or carried out in tandem, with industry. In the early 1980s, Kodak carried out a study with the Genealogical Society of Utah regarding the possibility of digitizing its millions of rolls of microfilm (see González 1992, 101), and IBM was involved in a variety of digitization projects, including those with the Archivo General de Indias (Spain), the lifetime collection of American painter Andrew Wyeth’s works (USA), the National Museum of Ethnology in Osaka (Japan), The Hebrew Union College (USA and Israel), Yale Library (USA), The Vatican Library (Italy) and the National Gallery of Art (USA) (Mintzer et al 1996, Gladney 1997). Such projects were beneficial to both institution and industry: providing the culture and heritage industry with access to expensive, advanced technology, and providing the industrial partners with real life test cases for research and development, publicity, and often financial benefits such as tax-breaks when working alongside charities, education, or governmental institutions.

However, international and national policy makers and funding bodies were taking notice: for example, “DigiCULT” – the European Commission’s programme for undertaking “EU-funded research on digital culture and digital libraries deal[ing] with leading-edge information and communication technologies for expanding access to and use of Europe’s rich cultural and scientific resources” (European Commission 2007) – emanated from interest in this early digitization work and was established in the late 1980s. Policy development and funding streams for digital libraries and heritage effectively began at European level at the same time: with the European parliament drawing attention to the importance of the emerging digital environment, and fostering technology-based, applied research programmes for libraries under the Third and Fourth Framework Programmes in the early 1990s (Manson 1998). By the close of the 1980s, funding bodies across the world, lead by the USA and the UK, were beginning to develop programmes to provide funds for institutions to undertake digitization projects.

The excitement in the humanities and heritage sector regarding the possibilities of digitization in the late 1980s and early 1990s was palpable, with much published research emanating from early pilot projects regarding the technicalities and possibilities of digitization. Thaller declared that “image processing in many ways has been the “hottest” topic in Humanities computing in recent years” (1992b, 1) after a series of workshops were held in Glasgow, Tromsø, and Firenze on the topic. The exponential growth of digitization in the arts, cultural and heritage sectors was heralded as

…a curious and unprecedented fusion of technology, imagination, necessity, philosophy and production which is continuously creating new images, many of which are changing the culture within which we live (Colson and Hall 1992, 3).

The introduction of the networked global information environment in the early 1990s, the resulting public policy encouraging the production of these resources, and the emergence of funding streams to create online resources, was to further encourage unprecedented investment and development in digitization within the library, heritage and cultural sector.
THE RISE OF DIGITIZATION

THE 1990s

The exponential growth of digitization projects being undertaken in the close of the 20th Century meant that "to the librarian at the very least, [the 1990s] could be termed the 'decade of digitisation'" (Lee 2002, 160). The increase in digitization efforts was due to a complex interplay between different forces: the increase in performance and availability of new networked technologies (Naughton 2000), an increase in awareness of the possibilities of these technologies. Additionally, the resulting changes in public policy which increased the availability of funding for memory institutions undertaking technological projects and provided infrastructure to facilitate digitization efforts, and the changing perceptions within memory institutions themselves regarding how technology could be appropriated to meet user needs and offer new possibilities increased digitization efforts further. This culminated in "Countless millions of pounds, dollars, francs and marks [being] ploughed into digital projects that have involved the conversion of library, museum and archive collections" (ibid).

In many ways, the 1990s began as the 1980s had ended, with small scale digitization projects being undertaken of treasured holdings. For example, The British Library’s earliest and most prominent project was that of the *Beowulf* manuscript, and shortly afterwards, the Bodleian Library began putting digital surrogates of its Medieval and Celtic manuscript material online. Early projects tended to build on and reinforce research efforts into best practice (the research involved with digitization rarely invented new technologies, but ascertained how technologies invented elsewhere for other purposes could be appropriated in memory institutions). Faced with limited computer storage and network speed, expensive capture technologies which had not yet reached maturity, and software which needed tweaking to fit the needs of memory institutions, these projects were pioneering experiments into what could be done, and how best to attempt it.

However, the development of the World Wide Web in the early 1990s (Naughton 2000), which provided a vehicle for dissemination of digital information, encouraged scoping studies and research to be undertaken to ascertain how the new imaging and networked technologies could be appropriated, such as Peter Robinson’s seminal study “The Digitization of Primary Textual Sources” (1993). Growing interest from the research councils themselves led to funding streams which encouraged an exponential growth in the creation of digital surrogates, on a worldwide scale, shifting the emphasis rapidly from small, in-house, projects of limited scope, to large scale, ambitious projects aiming to bring access to a high volume of resources to the masses.

In the early 1990s, following interest and resource provision by centralised sources, digitization of individual items began to gather pace. In 1993 in the UK, the Joint Funding Council’s Libraries Review Group published a report known as The Follet Report (after Sir Brian Follet who chaired the investigation) which was a major review of the needs of libraries in the changing digital environment (Joint Funding Council’s Libraries Review Group 2003). Following publication of the report, the UK government’s Joint Information System’s Committee (a centralised initiative funded by all four of the Higher Education funding bodies for Scotland, England, Northern Ireland, and Wales) announced funding of a string of initiatives
to encourage the development of Electronic Libraries (or ELib), related resources, and infrastructure and services to support the use of digital content in Higher Education and the institutions which serve it. More than £100 million was spent in this area by JISC by the close of the 1990s. In addition to this, JISC provided over £15 million of funding for the ELib programme (which looked into how digital libraries may be developed, including individual small scale studies into digitising particular content, see Wissenburg 2000 and Rusbridge 2001), £50 million on networks, and £24 million on data services (Rusbridge 1998). Other funding bodies were quick to follow, including the Leverhulme Trust, the British Academy, the National Heritage Fund and its related lottery counterpart the New Opportunities Fund, and the Arts and Humanities Research Board (AHRB). It is difficult to ascertain how much money in total went into digitization initiatives in the UK in the 1990s, but fair to say that it rapidly became an industry in itself, with most major academic, library, and museum institutions bidding for funding in an increasingly competitive funding environment.

To facilitate best practice in the creation of digital resources in the UK, and to provide long term storage facilities for resulting digital artefacts, it became clear that centralised services needed to be provided by the funders of digitization projects to ensure that standards were met, that the resources created were not lost upon the close of individual projects (Burnard and Short 1994), and that resources could be used by as wide a constituency as possible. In the library sector, the Consortium of Research Libraries in the British Isles (CURL, http://www.curl.ac.uk/) was founded in 1992, funded in part by JISC, to provide tools, infrastructure, and leadership, creating a cross-searchable library system to encourage easy access to digital resources by “helping to build the Distributed and Hybrid Research Library of the future” (CURL, n.d). The Arts and Humanities Data Service, formed to “collect, preserve and promote the electronic resources which result from research and teaching in the arts and humanities” (AHDS 2005) was founded in 1996, funded jointly by the JISC and the AHRB, incorporating the existing History Data Service and the Oxford Text Archive (AHDS 2006, http://ahds.ac.uk/index.htm). The Technical Advisory Service for Images, providing advice and guidance to the UK’s Further and Higher Education community regarding the creation of digital images, managing digitization projects, and delivering these to users for a variety of educational purposes (TASI 2002), was funded in 1997 by the JISC as a 3 year project, and given full status as a service in 2000 (http://www.tasi.ac.uk). These initiatives, and many others like them, served to provide the infrastructure to ensure that the quality of digital resources produced was high, and that the resulting digital resources would be made available to users.

In addition to creating digital images and surrogates of objects, and digital records of existing physical holdings, it was also necessary to investigate the infrastructure in which these records and digital representations could be hosted, searched, stored, and delivered to the users. The concept of the “Digital Library” as an all encompassing provider, store, and search mechanism emerged throughout the 1990s:

Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret,
distribute, preserve the integrity of, and ensure the persistence over time of
collections of digital works so that they are readily and economically available
for use by a defined community or set of communities (Digital Library
Federation 1998).

Much work was devoted to establishing digital library initiatives and providing
digital content for such endeavours from the early 1990s. However, it is still not
clear what constitutes a digital library:

The digital library, the electronic library (generally taken to be synonymous
with the digital library), the virtual library, the hybrid library, the library
without walls are all concepts that librarians seems to be dealing with all the
time. What do they mean? Do they mean the same to everyone who uses the
terms? Do they all mean the same thing? Do we all mean the same thing when
we talk about a library? … Every library is different, every digital library is
different, and different players are advancing many definitions for the digital
library (Deegan and Tanner 2002, 20).

Implementations of digital libraries from the 1990s onward explored the concept,
and the various possible interpretations of the term.

In the USA, the National Digital Library Program was initiated by the Library of
Congress in 1994, reaching its goal of making 5 million historical American items,
from over 90 collections, and various US institutions, available on its American
Memory Web site (http://www.loc.gov) by 2001. This public-private partnership
resulted in over $130 million being donated to the Library; from government,
businesses, and individuals (Library of Congress 2001). Another programme, the
Digital Library Initiative, was launched by the United States National Science
Foundation with the National Aeronautic and Space Administration (NASA) and
the Advanced Research Projects Agency (ARPA) in 1994, giving digital libraries
a major boost and allowing exploration of some of the major issues confronting
the establishment of digital libraries on a large scale. Phase 1, running from 1994
to 1998, funded six digital library projects with a total of $30 million. Phase 2
(DLI2, http://www.dli2.nsf.gov/) running from 1999, provided $55 million to fund
36 projects which aimed to extend and develop innovative digital library techno-
logies and applications. Clearly, a lot of activity was going into the creation of digital
resources, but also investigating the infrastructure necessary to provide them. Other
funding bodies in the USA also began to support digital library and digitization
initiatives during this phase, such as the Research Libraries Group (Erway 1996),
JSTOR (Schonfield 2003), The Getty Institute, the Andrew W. Mellon Foundation,
the Coalition for Networked Information, and the Art Museum Image Consortium
(AMICO) programme. Some of the funding bodies did not and do not limit their
support to the USA, but are active in pursuing international collaboration that
builds research collections without regard to their physical location.

The centrally funded infrastructure which underpinned these initiatives in the USA
mostly centred around the concept of the Digital Library. Digitization initiatives
followed long term attempts to share library catalogue information: founded in 1967,
the OCLC (Online Computer Library Center, http://www.oclc.org/) is a “computer
library service and research organization dedicated to the public purposes of furthering access to the world’s information and reducing information costs” (OCLC 2007) through the sharing of library records, providing services for locating, acquiring, cataloguing, lending and preserving library materials now used by over 60,000 institutions in 112 countries, and funded by a variety of international governments, institutions, and charities. The Coalition for Networked Information (http://www.cni.org/), founded in 1990 by the library and IT Communities to “enhance scholarship and intellectual productivity” (CNI 2006), rapidly grew to a current membership of 200 institutions representing higher education, publishing, network and telecommunications, information technology, and libraries and library organizations. The Digital Libraries Federation (http://www.diglib.org/), founded in 1995, comprises of a consortium of the main academic and public libraries, providing leadership and support for large scale digitization programmes, guidelines, and forums for sharing experiences between members. In addition to such consortia and forums, vehicles for disseminating research and reporting of best practice in digitization were established. D-Lib Magazine, an electronic publication with a focus on digital library research published by the D-Lib Alliance, launched its first issue in July 1995, and is still publishing related research and developments today for a wide international readership (http://www.dlib.org/). The Research Libraries Group (now part of the OCLC) started publishing their newsletter, Diginews, an online newsletter for digital imaging and preservation in April 1997, which ran for 10 years, and is now preserved at http://digitalarchive.oclc.org/. The founding (and funding) of such broad initiatives ensured that the communities undertaking digitization were sharing information and expertise, and working towards common goals.

A similar tale is to be told across most of the developed world. The European Union’s Telematics for Libraries Programme, running from 1990 to 1998, (http://cordis.europa.eu/libraries/) provided over €30 million to develop innovative library services and tools as well as bibliographic resources and library networking infrastructures which underpin such services, across Europe. A national programme was started in France in 1992 (Beagrie 2003), with the Bibliothèque Nationale de France undertaking an ambitious digitization programme of material in the national library and associated collections. By 2008 this ongoing digitization effort, one of the costliest and largest in Europe, had succeeded in making 10 billion documents available online (http://gallica.bnf.fr/). National reports for countries involved in the preservation of digital information, including historical information about digitization programmes, and covering Australia, France, the Netherlands, the United Kingdom, and beyond, can be found in Beagrie (2003). A summary report of the digitization efforts of all the EU member states has been produced ever year since 2002 by the Minerva project (MInisterial NEtwoRk for Valorising Activities in digitization, http://www.minervaeurope.org/), which also provides guidelines for best practice in digitization (Minerva 2007). DELOS, a Network of Excellence on Digital Libraries, provides guidance, advice, and forums to encourage the integration of efforts in Digital Library research and development across Europe (http://www.delos.info/), and its records also demonstrates the range and interest in digitization at this period.
The shift in digitization projects in the 1990s from the small scale and focussed, to large scale and all encompassing can be seen in the approach taken by one of the world’s largest and finest memory institutions: The British Library. Founded in 1972 by the amalgamation of various UK institutions, with a collection of over 150 million individual items gathered over 250 years which cover all ages of written civilisation, the British Library began to seriously consider digitization in the early 1990s. British Library’s digitization projects began with the *Electronic Beowulf* in 1992 (Kiernan 1997, Prescott 1998), a collaboration which

…linked access, scholarship and conservation of the fire damaged manuscript. Early concentration on material such as Beowulf which was unique, valuable, at risk (in conservation terms), and of high scholarly importance, was typical of the first steps by major libraries in the early 1990s (Smith 2006, 9).

Initial experience gained through the *Beowulf* project at the British Library “was built upon and improved through rapid advances in and general availability of technology” (ibid), which “saw an important shift in the mid-to-late 1990s from digitization as experiment to digitization as an operational task” (ibid). Since 2000, the BL’s digitization strategy has focussed on large scale digitization of textual materials to build a “critical mass” (ibid) of material, such as the major newspaper digitization projects: the digitization of a million pages of the Burney collection of 18th Century British Newspapers (http://www.bl.uk/collections/burney.html), and the digitization of two million pages of 19th Century local, regional, and national newspapers (http://www.bl.uk/collections/earlynewspapers.html). The library’s materials are increasingly now produced

…primarily – or solely – in digital form, and the Library’s digital collections are large and rapidly expanding…For the BL, digitisation offers an important way to open up its unique heritage collections, and make more widely accessible items of national and international appeal (Smith 2006, 5–7).

At the start of the 1990s, then, “In the early days of the Web, museums provided some of the best content and some of the most compelling reasons to go on-line” (Peacock et al 2004). Towards the close of the millennium, “A decade of digitisation and documentation for the Web … created a rich array of cultural and historical information across the museum, library and archive sectors” (ibid). The development of the World Wide Web was the cause of the growth of digitization efforts in the library, cultural, and heritage sector, and its increasing popularity with users fanned the expectation and desire to see digitised material online:

The internet has stimulated a demand for online access to information. Institutions, like museums and libraries, correctly perceive that digital images – at the same time as they reduce the need for direct access to the originals – can be used to provide improved access to those works in their collections that are considered most important (or most “in demand”). Importantly, such access is location- and time-independent. That is, scholars anywhere in the world can retrieve or view web pages (Grycz 2006, 33).
With the growth of material online in the latter part of the 1990s, it became almost impossible to track and trace all digitization projects. Efforts such as ITEM database switched from being a print publication in 1995 (Pring 1995) to being an online resource, but folded in 2000, given the impossibility of the attempts to keep cataloguing this type of online resources, and the success of the Internet in allowing individuals to search for the type of projects they wanted online. The act of finding relevant material, including digital image material, amongst the host of information available online became the subject of much scholarly effort itself (Terras 2008). Around the same time, the process of digitization, explored and experimented with by early projects, and ratified into procedural and technical guidelines along the way (see Besser and Trant 1994, Kenney and Chapman 1996, Kenney and Reiger 1998, Lee 2002, Deegan and Tanner 2002, Hughes 2004, MacDonald 2006), became less of a scholarly endeavour within itself, and more of a standard means to provide digital information to a wider audience. By the end of the decade,

Digitisation is not a per-se research issue but is part of a wider context related to the information society and the effective use of the digital content by cultural institutions (Mmerva 2003, xxiii).

POST 2000

Much digitization effort accompanied the hype surrounding the millennium celebrations, with many funding programmes in place to provide the resources to create digital content. These efforts continued into the new century: for example, in July 2001, the UK Government’s lottery based New Opportunities Fund (NOF) announced the largest single co-ordinated funding initiative in the UK for digitization, with grants totalling £50 million for more than 150 UK local and national organisations to document a wealth of UK achievement and heritage in science, culture, and social diversity under its NOF-digitise programme. Such a wide reaching programme was undertaken to create a “communities bank” of Internet resources, primarily for use in learning, and targeting, in particular, “lifelong learners”. For a summary of the scope of these projects, explore http://www.enrichuk.net/. The JISC also announced a large scale digitization programme in 2004 (http://www.jisc.ac.uk/digitisation_home.html), aiming to “change the world of authoritative e-resources through its investment in digitising content from some of the UK’s greatest collections” (JISC 2007). The first phase, with funding of £22 million,

...began in 2004 and included the 18th century parliamentary papers, Archival sound recordings, British Library 19th century newspapers, Medical journals backfiles, NewsFilm Online and Online historical population reports (ibid).

The second phase, with a further £12 million,

...features projects that will enable all users, regardless of location and time, online access to a range of authoritative digitised e-resources previously
difficult or impossible to access. Together, the projects represent a diversity of rich and vivid perspectives on the history, culture and landscape of the UK and beyond. They capture a wide variety of aspects of UK life, from Cabinet papers to First World War poetry, radio news to East End music hall, political cartoons to British borders, and in a wide range of media, including sound, film, images, journals, newspapers, maps, theses, pamphlets and cartoons (ibid).

Funding councils, then, are now aiming to create a knowledge base of high quality resources, from established projects, institutions, and repositories, under their watchful eye, demanding observance of strict technical guidelines with aims for reuse in national repositories. The majority of these resources are (but are not exclusively) digital image based.

Around the same time as the launch of large scale, centrally funded initiatives such as the JISC digitization programme came the emergence of proposals from large, commercial firms (namely Google and Microsoft) to digitise holdings of institutions wholesale. By avoiding the “cherry picking” tactics of the funding councils, they aim to digitise everything possible, with the view that searching mechanisms to allow intelligent and useful analysis of large scale digital collections will soon become sophisticated enough to allow users to find their potential needles in the digital haystacks. This is often carried out in co-operation with world leading, large institutions to allow large scale digitization of holdings that they could just not afford to do on their own. However, such digitization programmes are rife with copyright problems, and their restrictive legal conditions can often cause complaint and consternation, and:

…some libraries and researchers worry that if any one company comes to dominate the digital conversion of these works, it could exploit that dominance for commercial gain (Hafner 2007).

As a result of both the commercial and publicly funded efforts to provide high quality material online, it is now difficult to remember a time when institutions did not provide digitised representations of their holdings, and the results of years of investment of effort and resources by the cultural, heritage, commercial, and voluntary sector provide a rich, online environment for users to browse, analyse, and study:

Do you want to tread Shakespeare’s Hamlet in the first quarto edition of 1603? You have only to go to the British Library site and click on the heading “Treasures in Full”. Do you need to consult a Finnish journal for a certain date in 1805? Go to the University of Helsinki Library site and the appropriate issue will appear on your screen. To consult descriptions of monuments in Egypt and Nubia, click on the site of the Maison de L’Orient et de la Méditerranée. And so forth (Jeanneney 2007, 19).

Indeed,

The possibility of examining rare and unique objects outside the secure, climate-controlled environments of museums and archives liberates collections for study and enjoyment (Besser and Trant 1995, 7).
The shift in digitization projects to the large scale, and the growth in use and user expectations regarding the provision and quality of digitised material, results in a change in research questions that need to be asked by professionals working in the area:

We’re getting pretty good at digitizing material at scale…. If you look at many of the projects … you will see there are lots and lots of materials being digitized. Our museums, our libraries, our archives, our historical societies, are all running digitisation programs…. We have a wealth of experience and a large number of successful projects (not to mention some highly educational failures) to build upon. With the exception of relatively esoteric materials in specialized formats or that have some really unusual characteristics, this is not really research any more. Or to put it another way, the research questions are less about how to do it at all and more about how to optimize – how to do it more efficiently or effectively, how to be sure that you’ve chosen the most appropriate strategies and technologies (Lynch 2002).

At the time of writing, research questions remain about use and usefulness of digitised collections, how new and developing technologies, such as those touted as “Web 2.0” can be integrated into digital online collections to be of benefit to users, and how to integrate the vast swathes of online ephemera digitised by amateurs into our thinking about digitization, collections, and corpora (Terras 2008).

Digital imaging technologies have been readily adopted by memory institutions, and, following a period of experimentation, it is now commonplace for libraries, archives, museums, art galleries, and private collections to undertake the costly and time consuming effort of producing digital surrogates of their collection, usually for delivery online, to provide the accessibility and digital coverage of collections that users have now come to expect. There are interesting times ahead for the information professional in dealing with the wealth of digital information now available, keeping up with the changing nature of Internet technologies, and coping with resulting changes in user expectations and needs, to create digital libraries which can exploit the potential of this groundswell of digital – digitised – content.

NOTES

1 The majority of digitisation projects in the 1980s used videodisc (Philip’s LaserDisk) as the delivery media: a technology which was soon superseded by CD-ROM, and eventually the Internet. Videodiscs needed their own players – now unavailable: as a result the majority of the output of these early projects is now unreadable. The most high profile example of this was the British Broadcasting Corporation’s Domesday Project, where schools around the UK were asked to survey their areas to produce a database of how Britain looked to the British in 1986, 900 years after the original Domesday book was compiled (Finney 2007). The resulting database of statistics, maps, and photographs were saved onto videodiscs for posterity, but by 2000 the majority of videodisc players were defunct, and a lengthy and costly program of conversion and emulation was begun in 2002 by LongLife Data, taking 16 months to recover the videodisc data and re-create the original application (Darlington et al 2003). LongLife Data created a web version of Domesday 1986, which is now available at http://www.domesday1986.com/.
THE RISE OF DIGITIZATION

REFERENCES


González, P. (1992). The digital processing of images in archive and libraries, large scale international projects. In M. Thaller (Ed.), Images and manuscripts in historical computing, proceedings of
THE RISE OF DIGITIZATION


TERRAS


*Melissa M. Terras*

*Department of Information Studies*

*University College London, UK*
2. DIGITAL LIBRARIES AND DIGITISATION

An Overview and Critique

INTRODUCTION

Libraries have been, and continually are, being transformed. I have written a considerable amount about the value of the library, particularly the public library (see Rikowski, 2001a, b, 2002a, b, c, d, e, 2003b, 2005), but with the slant being more towards the traditional library. However, the digitisation programme has expanded rapidly, particularly over the last 10 years or so, and the traditional library has been and still is undergoing much change. The term the ‘hybrid library’, combining both the traditional library and the digital library, is being referred to more and more today, although obviously it is still not in common usage amongst the general public.

Digitisation offers many new and exciting opportunities, and indeed, it is changing our way of life in many ways. It certainly means that much more information is easily and readily available. Yet, there is no way that it can wholly replace the physical, traditional library, with the opportunities for people to engage in dialogue, and to build a sense of community. Neither can the e-book ever replace the hard copy book, in my view. You cannot take an e-book to bed with you!

This chapter will provide an overview and a critique of digitisation and digital libraries.

DEFINITION OF ‘DIGITAL LIBRARIES’

It is always important to define terms, so first of all, let us consider the definition of the ‘digital library’. Hughes says that digitisation is ‘the process by which analogue content is converted into a sequence of 1s and 0s and put into a binary code to be readable by a computer’ (Hughes, 2004, p. 4), whilst Lynch says that the ‘digital library’ can mean different things to different people and that:

The field of digital libraries has always been poorly-defined, a ‘discipline’ of amorphous borders and crossroads, but also of atavistic resonance and unreasonable inspiration. ‘Digital libraries’: this oxymoronic phrase has attracted dreamers and engineers, visionaries and entrepreneurs, a diversity of social scientists, lawyers, scientists and technicians. (Lynch, 2005, p. 1)

Lesk emphasises how building digital libraries will change our understanding of organisations, and perhaps even change our culture:

Building digital libraries is not just a question of piling up disk drives: it involves creating an entire organization of machines and people, perhaps even...
a culture, in which we are able to find information and use it. The social implications of a world in which information is distributed almost without institutions are not understood; what does this mean for universities, for education and for publishers? (Lesk, 2005, p. 5)

The ‘digital library’ is also referred to by a variety of other terms, such as the ‘electronic library’, the ‘virtual library’, the ‘library without walls’, the ‘online archive’ and the ‘desktop museum’. The term was originally adopted by the computing science community, whilst librarians, at the time, were still referring to the ‘electronic library’ and the ‘hybrid library’.

Many different types of material can be digitised, including pictures, theses, books, journals, newspaper articles and magazines. As Crane says:

The digital libraries of the early twenty-first century can be searched and their contents transmitted around the world. They can contain time-based media, images, quantitative data, and a far richer array of content than print, with visualization technologies blurring the boundaries between library and museum. (Crane, 2006, p. 1)

The scale of the increase in digitisation has been very rapid over the last few years. By mid 2003, for example, Google claimed to be indexing over 3 billion web pages, and the Web has over 150 terabytes of text.

Jones et al (2006) argue that there is no consensus in regard to the definition of a ‘digital library’. They also highlight the difference between the ‘institutional repository’ and the ‘digital library’, arguing that the ‘institutional repository’ exploits intellectual capital produced by an institution, whilst the ‘digital library’ is a broader collection, and also includes materials published elsewhere. Furthermore, they suggest that a collection of items is not a digital library just because the items have been digitised. Rather, it is a library first and it cannot just be a commercial project: ‘A true digital library has to be organised for its purpose, and must not be randomly heterogeneous and indexed as a commercial by-product’ (Jones et al, 2006, p. 4).

But to what extent is it possible, anyway, for non-commercial organisations to be largely responsible for digitising the information and the knowledge of the nation, as it were? Can they afford to do this? On the other hand, can information continue to be made freely available? Of course it can be, if we chose to put our resources into any such project, but in capitalism, that is just not a realistic possibility, in any all-encompassing way. This is because the drive to create value and profit will always override other considerations, including the importance of providing free access to information.

Hughes concludes by referring to digitisation in the following way:

Digitization of cultural heritage materials is changing the ways in which collections are used and accessed. Many materials are amenable to digitization, including scarce, fragile and ephemeral materials, as well as the whole spectrum of moving image and audio materials. All can be safely used by a wider audience in digital form. (Hughes, 2004, p. 29)
There are many advantages to be gained from digitisation (thereby aiding with the creation and expansion of digital libraries). Hughes usefully outlines a number of them. As Hughes indicates, with digitisation one can ‘search, browse and compare materials in useful and creative ways’ (2004, p. 10). Hence, there is a much greater and easier access to many different types of material and collections. E-books also can provide very good searching facilities, and they take up less storage space than hard copy books.

Some organisations, such as the National Gallery in London, even create an electronic image of every item in their collection. Or sometimes organisations collect electronic images around themes. One example here is the Arnamagnaeanae Institute at the University of Copenhagen, which is a project that has a web-accessible catalogue of medieval Icelandic manuscripts, with the aim being to achieve a ‘virtual reunification’ of the two halves of the Arnamagnaeanae collection. It is also possible to develop a digital surrogate of a rare or fragile original object. This means that people will have access to the object, helping to ensure that the original is not damaged by a lot of handling.

Sometimes, collections have to be split up, over long distances. This can be quite disappointing and inconvenient for the user. The Management Committee of the Feminist Library in London in 2007, for example, were discussing the possibility of having to break up the feminist collection at some point, if they were not able to attract enough funds to preserve the library intact. But through digitisation, whole collections can be brought together by a virtual means.

Other advantages of digitisation (as outlined by Hughes) include the fact that resources can be put to unforeseen uses in the future; that delivery can be satisfied immediately and that it can bring prestige to and raise the profile of an organisation. Also, detailed records can be created about the collections and staff can develop their knowledge and understanding of the new technologies. New ways of preserving, collecting, organising, accessing and disseminating knowledge and information can be developed with digitisation.

Digitisation provides educational benefits as well. John Unsworth (2000) argues that networked digital information can really encourage, support and enrich the essential elements of scholarship. Material that is digitised can be accessed and read in new, different and improved ways. Material can often be read quickly and easily, but obviously it also has to be undertaken with caution, so as to avoid possible health risks. E-learning is also becoming an important part of education, and the notion of the ‘virtual campus’ is growing. An example of how digitisation is changing the essentials of scholarship can be seen through the development of an online edition of Chopin’s 1st Editions Project, which is based at the Royal Holloway College, University of London. The online edition is being used to analyse the creative history of Chopin’s music.

Digitisation is also an important means for preserving our cultural heritage as well as enabling advanced research to be conducted on historical materials. This includes digitising ancient manuscripts (see, for example, Nikolova-Houston & Houston’s chapter in this book). However, as Beagrie (2004) says, there is a serious
gap between our ability to create digital information and our infrastructure and capacity to manage and preserve it over time. This is sometimes referred to as the ‘digital dark ages’. Furthermore, decisions have to be made about what historical material to digitise and what not to digitise. As Hughes notes:

…most institutions will integrate many different technology-based projects over a long period of time. Some of these projects will overlap, some may ultimately contribute to an institutional ‘digital library’, while others may become known as ‘legacy projects’, leaving preservation concerns and headaches for future caretakers. (Hughes, 2004, p. 17)

Overall, then, there are clearly many benefits to be gained from digitisation, a few of which have been outlined here.

COSTS OF DIGITISATION

However, digitisation is also costly and there are indirect as well as direct costs. Certainly, it is not easy to make short-term cost savings. Many questions can be raised here. How concerned do we want to be with getting good returns from our expenditure on digitisation, for example? And how can we measure this anyway? On the other hand, can digitisation cause information to ‘lose’ some of its value? Lesk (2005), for example, asks us to consider what the value of information is, and who is likely to be willing to pay to preserve digitised information. Yet, as he points out, one cannot put a numerical value on delivering information to the desktop instead of within the traditional library.

Hughes also notes the importance of focusing more on long-term strategies for digitisation, particularly given the fact that budgets are limited. She says that:

In recent years, a growing understanding of the costs of digitization, in terms of both time and financial resources, has placed a greater focus on developing digitization initiatives and programmes that will realize tangible and strategic benefits for the institution and its users, rather than opportunistic or short term projects that are limited in their scope or focus. (Hughes, 2004, p. 8)

A number of studies have been undertaken exploring the likely costs of digitisation. Shelby Sanett (2002), for example, usefully proposed the development of a cost model for preserving authentic electronic records. Her cost model broke costs down according to capital costs, direct operating costs and indirect operating costs (overhead) for both preservation and use of electronic records. Costs included labour, space, materials and equipment. Mackie-Mason et al (1999) examined different pricing and bundling models of electronic resources in relation to usage and publisher revenues.

Connaway & Lawrence (2003) compared library resource allocation for paper with the digital library. This was an exploratory study that asked 11 Association of Research Libraries (ARL) librarians to identify resources needed in transforming from an all-paper to an all-digital library. Within this, they looked at the selection, acquisition, organisation and dissemination of paper and electronic materials. It was assumed that fewer labour, space and material resources would be needed in an
all-digital library, but that equipment resources would be the same. Labour included professional, paraprofessional and technical staff. Space included work, private and public space, whilst equipment included bookshelves, tables, network connectivity equipment, terminals and printers, etc. Librarians had to consider two hypothetical types of libraries; a ‘paper’ library consisting only of paper books and a ‘digital’ library consisting only of electronic books. It was concluded that a mixed paper-digital library would probably cost more than having either a completely traditional or a completely digital library. Digitisation might increase costs in the shorter term and it might be some while before the full cost benefits of digital media were fully realised. Yet, the findings suggested that the costs for labour, space and material were likely to be less in an all-digital library compared to an all-paper library.

In sum, quite a lot of useful studies have been undertaken investigating the costs of digital libraries, a few of which have been considered in this section. People and organisations can use such studies, and the cost models derived from them, to help them to decide whether to digitise their stock and if so, which parts of it to digitise, and how much of it to digitise.

JISC

JISC is an important body in the United Kingdom (UK) in regard to promoting and developing digitisation. The UK Joint Information Systems Committee (JISC) is a committee of UK Higher and Further Education Funding Councils and serves about 200 higher education (HE) institutions and over 500 further education (FE) institutions in the UK. Its aim is to promote innovative use of information and communications technology (ICT) in tertiary education.

JISC, Digitisation and Preservation Issues

JISC has had a considerable input into various digitisation projects and digital preservation over the last few years (see Beagrie, 2004). It plays several roles in regard to digital preservation. This includes supporting collective action for the sector and advising institutions. It also includes finding appropriate licensing arrangements for commercial content and building long-term preservation with other agencies. If digital preservation issues are not properly addressed, our cultural heritage could well be under threat. JISC has been developing different toolkits to help institutions with their information-related issues.

In 2000 Beagrie was appointed by JISC to build on the digital preservation work that it had been undertaking. There were three major objectives of the project. The first was to establish and disseminate best practice and guidelines for digital preservation; the second was to encourage different organisations to work together, on digital preservation issues, in order to generate funding on a global basis; and the third was to develop a long-term digital preservation strategy for the FE and HE sector in the UK.

Beagrie emphasises that digital preservation needs to be seen as a means to an end and not as an end in itself: ‘Ultimately, digital preservation will be successful
when it can be seen not as a stand alone institutional activity but as an activity embedded in how institutions manage and approach digital information and resources on an ongoing basis’ (2004, p. 11).

The ‘JISC Digitisation Programme Blog’ started in April 2006. It highlights significant e-resources from the UK. Following on from this, also in April 2006, JISC called for proposals for digitisation. Proposals submitted included the digitisation of Icelandic manuscripts; all the photographic negatives held by the Scott Polar Research Institute; the Carl Giles newspaper cartoon archive; and rare pamphlets and newspapers from the Anglo-Jewish community; the Bristol University Fossil Record Database; and King’s College London: Gramophone Magazine.

In sum, the work that JISC has undertaken and continues to undertake in regard to digitisation and digital preservation is very important and continues to grow.

TRADITIONAL, DIGITAL, HYBRID AND VIRTUAL LIBRARIES

In regard to traditional libraries, Lagoze et al indicate that they:

…are much more than well-organized warehouses of books, maps, serials, etc. In their full expression, they are places where people meet to access, share, and exchange knowledge. The resources they select and services they offer should reflect the character of the communities they serve. (Lagoze et al, 2005, p. 2)

Borgman suggests (1999, 2000, 2003) that digital libraries should match and then extend traditional libraries. However, the digital library can surely never completely match and replace the traditional library. Meeting and sharing knowledge virtually, is not the same, and surely never can be the same, as meeting face to face. Lagoze et al point out that digital libraries:

Like any library … should feature a high degree of selection of resources that meet criteria relevant to their mission, and they should provide services, including search, that facilitate use of the resources by their target community. But freed of the constraints of physical space and media, digital libraries can be more adaptive and reflective of the communities they serve. (Lagoze et al, 2005, p. 2)

Hence, in this way, they should be able to cater well for the wants and needs of communities. Lagoze et al also say that digital libraries should be collaborative, and enable users to contribute knowledge to the library. Furthermore, they should be contextual and express ‘the expanding web of inter-relationships and layers of knowledge that extend among selected primary resources’ (Lagoze et al, 2005, p. 2).

Digital libraries can sometimes seem to be a catalogue of metadata records. Yet, as Reeves says, digital libraries are also cognitive tools and not just repositories of information:

The real power of media and technology to improve education may only be realized when students actively use them as cognitive tools rather than simply
perceive and interact with them as tutors or repositories of information. (Reeves, 1998)

There are various issues that need to be considered in the future. Will the hybrid library (combining both the traditional library and the digital library and which is also referred to in both Pandian’s and Dunlap’s chapters in this book) develop much further, for example? Or will the digital library increasingly start to replace the traditional library? How will people adapt to all this? And what implications will all of this have for society at large? And now we have a new library ‘on the scene’ – the virtual library (referred to in the chapters by Bell et al and Peters in this book). How will this develop in the future? Will it become a significant part of the library world? Bell et al have clear views on this and they ‘believe that libraries do need a presence in virtual spaces where more and more people are gathering’ (Bell et al).

THE LIBRARY COMMUNITY AND DIGITISATION

What about the response of the library community, in particular, to digitisation? Pearson argues that there is no shortage of action here, but that, on the other hand, there is no real overall sense of direction. Pearson says:

I believe there is an urgent need for the library sector to pull its act together and for us to find a mechanism to put digitisation of the documentary heritage, and a strategy for achieving it, high on the agenda. Making it all happen is not something that libraries can achieve single-handedly but inertia will lead to regret in due course. It is one of the great visionary challenges for the present professional generation. (Pearson, 2001, p. 5)

We also need to consider to what extent digitisation is changing the role of the traditional librarian. Hughes (2004) asks whether libraries and other organisations that are digitising content and making it available to many are becoming more like publishers, for example. If so, how will this impact on the work of the traditional librarian?

Lynch (2005) makes the point that some think that digital libraries have little to do with the practice of librarianship. Digitisation actually comes from a technical and engineering basis, from the 1960s. Lynch argues that we can see digital libraries as offering a ‘mature set of tools, engineering approaches, and technologies that are now ready to be harnessed in the service of many organizations and many purposes’ (2005, p. 4).

On the other hand, if digitisation is largely the responsibility of librarians and information professionals and they do hold the purse strings, then this raises other questions. Pearson says that various assumptions are made about digitisation in this regard. One is that the technology is now well developed and mature and is worth investing money in. And the other is that libraries will be funded in the same way that they always have been and that free access to documents will still be seen to be worthwhile. But can library budgets be stretched this far and will finance mechanisms in reality remain the same? And to what extent is it possible to have
free access to substantial amounts of digital information in capitalism? The cost models outlined above should provide some help and guidance to librarians in this regard.

Various research has been undertaken in regard to digital libraries. An important Digital Libraries initiative took place with organisations such as the National Science Foundation, the National Library of Medicine and the Library of Congress in this regard. In a document outlining Phase 2 of the project (in 1998), it said that the Initiative:

…looks to stimulate partnering arrangements necessary to create next-generation operational systems in such areas as education, engineering and design, earth and space sciences, biosciences, geography, economics and the arts and humanities. (Digital Libraries Initiative – Phase 2, 1998, p. 2)

Three specific research areas were being explored: firstly, human-centred research, which ‘seeks to further understanding of the impacts and potential of digital libraries to enhance human activities in creating, seeking, and using information and to promote technical research designed to achieve this’ (Digital Libraries Initiative – Phase 2, 1998, p. 3); secondly, content and collections-based research; and thirdly, systems-centred research, which focused on the integration of technologies, in order to develop flexible information environments. Clearly, this type of research is important and needs to be developed further in the future.

Obviously, digitisation impacts on many different types of librarians and libraries. Murray (2000), for example, considers the expanding role of the school librarian. Traditionally, the role has included teacher and instructional partners, she says, whereas the new role includes navigator, teacher and collaborator, evaluator, publisher, program administrator and staff developer.

The school librarian will continually be involved with new technologies, including digitisation (as indeed are librarians in general). As Murray says:

As information access becomes increasingly computerized, the school librarian will be responsible for introducing new technologies to her teaching colleagues and students, enabling and empowering them to adopt information-literacy strategies that will make them independent lifelong learners. (2000, p. 2)

In general, journal articles are obviously of crucial importance for obtaining up-to-date information about digital libraries and digitisation. However, books are also valuable. *Digital Libraries* by Lucy Tedd & Andrew Large (2005), for example, introduces readers to the principles underlying digital libraries, and illustrates the principles by reference to a wide range of digital library practices throughout the world. They also look at the origin of digital libraries and why they have emerged and developed. *Understanding Digital Libraries* by Michael Lesk (2005) includes chapters on topics such as the evolution of libraries, images and knowledge representation schemes. The first few chapters of the book examine how the contents of a digital library are stored and organised. There is a chapter on ‘multimedia storage and retrieval’, for example, which examines digitisation of sound, video, software, pictures and automatic speech recognition. Lesk also looks
at different digital library projects around the world, noting, for example, that US
digital library projects have largely been undertaken by the National Science
Foundation.

In sum, this section has outlined a few of the key issues in regard to digitisation
for the library community. There are clearly many other areas that also need to be
considered. Looking to the future though, Lynch says that:

The next decade for digital libraries may well be characterized most pro-
foundly by the transition from technologies and prototypes to the ubiquitous,
immersive, and pervasive information and information technology landscape.
(2005, p. 6)

E-BOOKS

Gunter says that ‘An e-book, in the broadest sense, comprises any book or
Many different types of material are being digitised, including e-books and e-theses
(see, for example, Susan Copeland’s chapter in this book on electronic theses).
This digital material can be made available in the digital library. This section
focuses on e-books, by way of illustration, and considers some research that has
been undertaken on e-books, the impact of e-books and how this is likely to
develop in the future. The growth in e-books over the last few years has certainly
been significant.

John Cox (2004) examined e-books, and described a case study from 2001,
where Librarians of the Conference of Heads of Irish Universities established a
working group to assess the e-book market and its potential for university libraries.
It was found (in 2002) that the market was volatile, and that there were problems,
in particular, with regard to poor on-screen presentation. It was also found that
e-books could support learning well in subjects such as business and law, where
information is structured in quite clear and discrete blocks.

As part of the study, the group recommended a one-year subscription to an
e-books service, focusing on business and computing. It chose Safari Tech Books
Online, which is owned by Pearson Education and O’Reilly. This has the full text
of 2000 titles and subscribers include over 400 academic libraries. Safari proved to
be very popular with users at the Irish universities that were being researched. It
was found that where content matched need e-books could support the academic
goals of the universities. Thus, the future for e-books was bright, especially if on-
screen reading became easier and customer-oriented screens and licensing could be
realised.

Gunter (2005) reported on a study of the early market for e-books in the UK.
Weekly web surveys were conducted with members of a large online panel, which
was established by a leading commercial Internet research company. Questions
explored subject areas such as awareness, trialling, purchase and borrowing e-books,
frequency of behaviour and types of publications accessed and/or obtained. There
were 3916 responses in total. From the survey it was found that 85% of respondents
were aware of e-books. Amongst these, about half had made trial use of them,
38% had bought at least one e-book, and 13% had borrowed an e-book from a library. The most frequently accessed e-books were technical books (42%), followed by popular non-fiction books that related to a hobby (32%). The less frequently accessed e-books were specialised research monographs (13%), and popular non-fiction – autobiography/biography (10%). Respondents gave three main motivations for buying e-books: convenience, they were cheaper than conventional books and it was easier to search the contents. Other reasons given included the fact that e-books were easier to store, they offered a multimedia format, were easier to carry around, they could be electronically annotated, and it is easier to make copies of e-books for friends. Early e-book users saw electronic reading largely as being for reference work and 56% preferred not to read long passages on screen. In the study, it was found that, in general, e-book use is fairly specific in type. It concluded that, perhaps, e-publishers should focus more on reference books in the near future.

There have been different and varying predictions about the likely growth of e-books but it is safer to rely on real sales figures. According to Macworld (2004), e-book publishing is growing faster than any other form of publishing. There has been a particularly significant growth in e-books in the USA. The Association of American Publishers highlighted the fact that the e-book segment grew from $211,000 in net sales in January 2002 to just over $3.3 million in January 2003. This represented a 1,447% growth rate, which was the biggest year-on-year increase of any category of publishing (Blough, 2003).

In conclusion, the e-book market and e-book publishing is growing and many new publishing companies are being established that produce e-books and sell and distribute them over the Internet. From various studies that have been undertaken with e-book readers, a number of advantages to e-book reading have been identified; these include convenience, lower cost and being easier to search. E-publishing can, on the whole, move much faster than traditional publishing and it is likely to be an increasingly important part of publishing in the future. E-books certainly form an important part of the whole digitisation programme.

A SELECTION OF DIGITAL LIBRARY PROJECTS ON A GLOBAL BASIS

Many different digital library projects are taking place and a few of them are highlighted in this section. Google, in particular, has been very proactive in regard to taking forward various digital library projects. As Crane (2006) notes, Google has created a consortium of libraries to build a huge digital collection on a global basis and there seems to be a lot of money available to convert books into digital form – Google is in the lead here, with Open Content Alliance (see below) following close behind. As Crane points out though, there have also been various barriers, such as a need for compulsory licensing.

The library of the University of Wisconsin-Madison, USA is one example of a library that has joined the Google project (see: Book Digitisation Project Gathers Pace, 2007). The library will provide access to hundreds of thousands of public and historical books and documents from over 7.2 million holdings at University of Wisconsin-Madison Libraries and Wisconsin-Historical Society Library. The project
initially focused on library collections that are free of copyright restrictions. Furthermore, it will target other collections that are used a lot, such as the history of medicine and engineering and patents and discoveries.

Part of the library at the University of Oxford is also being digitised (see Oxford-Google Mass-Digitisation Programme, 2006).

For most of the Bodleian’s long history, of course, the accessibility of the Library’s collections has been almost entirely dependent on the ability of its users to come physically to Oxford. But the emergence of the Internet, and the scope for creating digital surrogates of library materials for networked availability, have radically altered the paradigm for access to the Library, opening up a whole new meaning for the Bodleian as a ‘library for the world’ in the 21st century. (Oxford-Google Mass-Digitisation Programme, 2006, p. 1)

The Oxford Digital Library began five years ago, with the help of the Mellon Foundation. In January 2003, after consultation with Google’s senior executives, work began to find a way of making the Oxford material available electronically, so that it could benefit both Google and the Bodleian Library. It was decided, first of all, to digitise printed books that had no copyright restrictions. Printed books were chosen rather than documents such as manuscripts. This was because they wanted to aim for critical mass as they thought this was the best way for making inroads into the vast collections, and secondly, because arrangements were already in place for tackling material not thought to be appropriate for mass-digitisation procedures. Much optimism surrounded the project and they saw themselves as being:

…part of a ‘giant leap for mankind’, in helping to ‘bring the world’s information to the world’. (Oxford-Google Mass-Digitisation Programme, 2006, p. 3)

Other libraries working with Google to digitise parts of their collections include the New York Public Library, the University of California and Madrid’s Complutense University, which is the largest university library in Spain.

Then, there is the mass-digitisation project – the ‘Open Content Alliance’ (OCA), which was announced in 2005, and led by Brewster Kahle (Young, 2006). Selected out-of-copyright books were being taken from libraries around the world and turned into e-books. This had the backing of Yahoo and Microsoft. Libraries have been working on their own on digital projects, but this project brings together digital content created by many different academic libraries. A specialised document scanner, developed by Internet Archive, called the Scribe, was used, to keep the price of scanning cheap.

The disadvantage of Google compared with Open Content Alliance is that one cannot easily download the data or print the whole book. In regard to OCA, Google said that:

We welcome efforts to make information accessible to the world. The OCA is focused on collecting out-of-copyright works which constitute a minority
of the world’s books – a valuable minority, but certainly not complete. (Cited in Young, 2006, p. 5)

Meanwhile, AlouetteCanada is a Canadian national digital online project, which brings together various digitisation projects across Canada and makes them accessible to the public (see Carlson, 2006). It is a portal site which can be demonstrated at different colleges and museums focusing on digitised materials about or of interest to Canadians. The institutions involved in the project include the Universities of Alberta, New Brunswick and Toronto. A project like this can be very useful, particularly given the fact that Canada is very large geographically, but has such a relatively small population. Most of the money for the project came from public funds.

Ernie Ingles, Chief Librarian at the University of Alberta, compared the project with the privately funded and secretive Google digitisation projects at institutions such as the University of Michigan, saying that:

Google is making content available freely, but it is making that content available in a commercial way. The question is, will Google always be there in perpetuity to make that content available? (Ingles, cited in Carlson, 2006)

Another digitisation project is the Glasgow Emblem Digitisation Project (2004). The Glasgow Centre for Emblem Studies was awarded a grant of £163,385 by the Arts and Humanities Research Board for digitisation of the corpus of French sixteenth century emblem books under the Resource Enhancement Scheme. The aim was to have a sophisticated website, including high-quality images of approximately 5500 pages, fully searchable text, and full indexes for both text and images.

The GATT Digital Library: 1947–1994 provides access to documents and information about the General Agreement on Tariffs and Trade (GATT), which promoted trade barriers amongst its member states from 1947 to 1994. The digital library was part of a cooperative project between Stanford University Libraries and Academic Information Resources and the World Trade Organisation. A grant for the project was obtained from the Institute for Museum and Library Services, which is an independent US federal agency supporting a variety of preservation and access programmes. The library has over 30,000 public documents, and 300 publications of GATT are accessible from the site.

There are many other interesting and important digital library projects taking place. Examples include the University of Central England Electronic Library; the Networked Digital Library of Theses and Dissertations; the National Science, Technology, Engineering and Mathematics Digital Library Program; Variations projects for music at Indiana University; and the Library of Congress digital programme. Other national digital programmes are also growing, such as Denmark’s Electronic Research Library, which is creating a portal for Danish research libraries.

The concept of the ‘global digital library’ is also being referred to today. Having free access to digitised information globally is a wonderful idea, although somewhat unrealistic! Petrelli (2006) explores some of the developments and projects around the notion of the ‘global digital library’.
One such project is the *Bricks Project* (which can be viewed at: http://www.brickscommunity.org/prj). This is a European 6th Framework-funded project, which aims to integrate existing digital resources (i.e. digital museums, digital archives) into a common digital library, operating as a European collective memory. There is an open collection and the distribution of resources is transparent to users, who access a single source of information.

‘Diligent’ (Digital Library Infrastructure on Grid Enabled Technology) is another European Union-funded project, which is run by Professor Neuhold. It aims to create a distributed digital library enabling access to distributed digital resources through the use of Grid computing. Professor Neuhold thinks that the digital library community must take a lead on important issues, such as long-term preservation, communities working collaboratively to create and maintain the global digital library and specifying reward systems. As Petrelli says:

> The optimists claimed the Grid makes it possible to create a European-wide multimedia digital library that everyone can access from anywhere in a seamless manner. The Grid supports better access to multimedia (e.g. playing video) of the highest quality, and can virtually integrate material that is located in different countries. (2006, p. 23)

In general, there has been a lot of experimentation and research into digital library development around the globe in the last 15 or so years.

Thus, there are many different and important digital library projects taking place globally, and a few examples have been highlighted here. Clearly, these projects are growing in number. One key question here is the extent to which the commercial digitisation projects are likely to grow in the future.

**GENDER ISSUES, I.T. AND DIGITISATION**

It is a well-known fact that the computer industry in general is very male-dominated (I have considered this elsewhere – see Rikowski, 2003c). So, what might the implications be for digitisation?

As Leslie Regan Shade succinctly says in regard to female computer scientists, ‘The statistics for women in the computer science field are dismal, revealing that only a small percentage of computer scientists and computer professionals are female’ (Shade, 1993, p. 2). Butcher also notes that ‘You don’t have to look hard to realise there are far fewer women in the IT industry than men’ (Butcher, 2003, p. 6).

Margolis & Fisher (2002) undertook some detailed research on women in computing. They conducted over 230 interviews with over 100 male and female computer science students (equally divided between the genders) over four years (from 1995 to 1999) at Carnegie Mellon University, USA. Multiple interviews were conducted asking students a variety of questions, such as their early experiences with computers, their home and school environments and their decisions to study computer science. They found that:

> Very early in life, computing is claimed as male territory. At each step from early childhood through college, computing is both actively claimed as ‘guy
stuff’ by boys and men and passively ceded by girls and women. The claiming is largely the work of a culture and society that links interests and success with computers to boys and men. (Margolis & Fisher, 2002, p. 4)

Furthermore, many of the women who studied computing at degree level lost confidence through their period of study, and started to doubt their own ability. Thus, women are clearly at a serious disadvantage, in many ways, in this IT age.

In regard to computer software specifically, which is obviously of particular importance for digitisation and the reading of e-material, most software has been programmed and designed by men, which often does not cater adequately for women. Frenkel, for example, talking about computer software in schools, says:

Any computer science curriculum, whether implemented in a wealthy or disadvantaged school must involve the selection of software. But studies show sex bias in educational software ... Children using software designed for the opposite sex are more anxious after they interact with the program, and that anxiety leads to lowered scores in the subject the program was intended to teach. (1990, pp. 5–6)

Sylge also discusses computer design and emphasises how most of it is undertaken by men, and that programming is a very male-dominated area. Furthermore, Woodward (2001) points out that men account for 79% of computer analysts and programmers. Sylge (1995) notes that this might be disadvantageous for women, as women and men do tend to think differently.

Margolis & Fisher also focus on this subject, emphasising, in particular, how boys tend to invent things (in this case, inventing and designing a computer program), whilst girls just use things that the boys have invented. Once again, girls are socialised into playing this type of role. They say that:

Today, the world of cyberspace is shaping our environment and our culture. Very little is unaffected by the onslaught of technology. The actual products of computer science change the way we do business, the interaction we expect from work, life, and pleasure, and the way we regard entertainment. If boys invent things, and girls use things boys invent, a cyberspace culture will inevitably reflect the desires and sensibilities of males to the exclusion and often denigration of females. (Margolis & Fisher, 2002, p. 12)

If females start to feel excluded in this way, then this could clearly have very serious consequences. We could find ourselves in a situation where females are focusing more on reading hard copy material, whilst males focus more on reading e-material. Males will then have more information available at their disposal and they will be able to access it all more easily. Information and knowledge brings power, so men are then likely to have even more power and privilege. So, if we are not careful, digitisation could increase the inequalities between the sexes still further. Frenkel warns us about the dangers of women being left behind, saying:

…what are the repercussions to our increasingly computer-oriented society, if women – about half the population and professional workforce – are not as prepared in this discipline as are men? (1990, p. 10)
Thus, there is a need for more women to be designing screens and layouts, so that their wants and needs are more adequately reflected, and women need to be more proactive in the IT world in general, I would suggest.

Simon (2006) presents the findings of some empirical research on women and technology which took place in 2001–02 in the UK. The aim of the research was to investigate women’s perceptions of technological change in general (and not just related to the library environment specifically), and it used the concept of the information society as a flexible framework. In-depth interviews and focus groups were conducted, grounded theory was used and it was very much a ‘grass roots’ view. Information was collected from 50 women, aged from 20 to 70+, from a range of backgrounds. Three main areas were addressed in the interviews: the concept of the ‘information rich and information poor’; temporal and spatial impacts of the new technologies; and the future of the community in the network age. In general, many of the women interviewed were very positive about the new technologies, although some also expressed doubts and concerns. Their main experience of the Internet had been obtained from a work-based experience. There was also a real sense of concern about being left out, or at least, left behind, in the IT age, although, some were not happy about the need to retrain.

Women account for 70–80% of the workforce in UK and USA libraries, and obviously, technology is becoming an important part of libraries today. And so, as Simon emphasises, women need to be very much a part of this new IT age:

...although analyses vary, there is little questioning of the importance of ICT in everyday life, and as such it becomes ever more vital to understand how women think about, and deal with, technology in their everyday lives.

(Simon, 2006, p. 478)

Opinions differ as to whether IT will open up new and exciting possibilities for females and overcome some of the barriers, or whether it is and will continue to be just another form of expressing and asserting male domination. Shade & We, for example, highlight many of the problems that need to be overcome but are optimistic that women can play a significant role in cyberspace. They say that we need to make:

...the Internet easily accessible to all people; making networking an attractive communications tool for women, by creating tangible and viable information and resources; and by encouraging young girls and women to become involved in the development and deployment of the technology. It also means creating a friendly online environment, one that allows women to speak their thoughts without having to hide their gender. The world of cyberspace is one which is being shaped daily by the millions of interactions on it, and women can contribute much to these exchanges. (Shade & We in Shade, 1993, p. 9)

In conclusion, I would suggest that women should be shaping the new technologies in general, and digitisation projects in particular, rather than letting the new technologies shape them.
There are many social and political implications leading on from digitisation projects and the creation and extension of digital libraries. This includes issues around the digital divide; the exploitation of workers in the developing world; the possible loss of a sense of community; the concept of ‘self’ in the IT age; the possibility that less physical space will be allocated to libraries and information resources in the future; and potential health risks through over-use of e-reading (backache, eye strain, etc.). There are also issues related to the likelihood of less face-to-face teaching and less contact with university tutors (due to the increase in e-learning, easy access to digitised information, etc.) and those related to the ever-changing nature of technology and the fact that technology often becomes quickly outdated. Some of these issues will be considered briefly in this section, but for me, what is important is to maintain an appreciation of the value of Marxism throughout all such deliberations.

Helm (2006) alerts us to the fact that a lot of the digitisation projects are achieved by the exploitation of labourers in the developing world. He says that a new category of work is emerging – the ‘digital factory job’. Thousands of people are inputting data and ‘These digital bricklayers are in a sense building the new information pyramid’ (Helm, 2006, p. 1). Helm reports that in Madras, India, for example, ‘editors’ making a fifth of US pay work extremely long hours in order to digitize archived American newspapers from the 1700s to the 1980s. In Boston, New York and Palo Alto, California, Google book workers manually turn each page of millions of library books, so that they can be scanned and made available to people that visit the Google website. Helm says that:

Such menial work with data and information is hardly new. But the growing fruits of such operations – gaining online access to historical papers, for example – have never been so close to the fingertips of the average person. (2006, p. 2)

He also emphasises that ‘there are plenty of people willing to work behind the digital curtain one keystroke at a time’ (p. 3).

Thus, whilst many of us might reap the benefits of digitisation, many of these benefits have been gained by the exploitation of workers, particularly of those in the developing world. This, then, creates even further divisions between the rich and the poor.

Then, of course, there is the digital divide itself, with the world being divided between those that have access to a computer and those that do not. Clearly, digitisation could make this division worse. Large international companies in the developed world, in particular, can gain huge benefits from free and easy access to digitised information. Meanwhile, for those of us that are living in the digital age, Zizek (2006) asks whether we are actually living in an age of digital democracy or whether it is more a ‘new tyranny of cyberspace’. He says:

Marxists and other critically-disposed thinkers like to point out how cyber-space equality is of course deceptive – it ignores the material disparities (wealth, social position, power or lack of it and so on). (Zizek, 2006, p. 30)
DIGITAL LIBRARIES AND DIGITISATION

So, how ‘liberated’ are we by this new IT age? Many talk about being feeling overwhelmed by the amount of emails they receive, for example, as well as feeling inundated by the sheer quantity of information that is ‘out there’ in cyberspace. Also, how ‘empowered’ are knowledge workers, in reality? (see, for example, Rikowskï, 2004a).

The concept of ‘self’ can also take on strange forms in this new digital age, argues Zizek. He says that:

The fact that I perceive my virtual self-image as mere play thus allows me to suspend the usual hindrances which prevent me from realising my ‘dark half’ in real life. My electronic id is given wing. (Zizek, 2006, p. 30)

He also makes the point that he never really knows his cyberspace partners and, equally, they never really know him!

Also, the fruits of our intellectual labour are being captured, stored and repackaged more and more, which can lead to further exploitation of labour, and this clearly has implications for society. Crane (2006) makes the point that ‘Digital libraries, where books read one another in however a rudimentary fashion, have already begun to separate intelligence and action from the human brain’ (2006, p. 4).

Petrelli considers the broader implications of digitisation, saying that ‘The digital age is changing not only the way we work, but how we think of ourselves and communicate’ (2006, p. 22). Crane also notes that:

…we also must consider the consequences not only for digital libraries but also for the intellectual life of the human race as a whole if the record of human experience becomes, in substantial measure, freely available online anytime and anywhere. (2006, p. 3)

However, the most important question to ask is ‘what is actually driving this digitisation project forward?’ Is it being driven forward by altruistic reasons, that digitisation will bring real and lasting benefits to humankind in general? Well, I do not think that anyone is that naïve! No, instead, it is, once again, the drives embedded within global capitalism itself that are driving this whole programme forward so vigorously. This relates to the theory (from an Open Marxist perspective) that I am developing in my published works (see, for example, Rikowskï, 2003a, 2004a, b, 2005, 2007).

Thus, to put it simply, we are now in the knowledge revolution, this being the latest phase of capitalism. As I say in my article, ‘Value – the life blood of capitalism’: ‘in the industrialised/developed world today we are entering into the knowledge revolution, where knowledge is seen to be the key to success’ (Rikowskï, 2003a, p. 161) and ‘value is the life-blood of capitalism. There is no resting place for capital – it seeks out new value everywhere and anywhere and this value is created by labour’ (p. 175).

This knowledge revolution depends on the exploitation of intellectual labour and the:

…creation of value from intellectual labour, which is then embedded in the commodity becomes necessary, so that intangible commodities can be sold in
the market-place and profits can be made (and ultimately profits can only ever be derived from value). (Rikowski, 2004b, p. 7)

Digitisation programmes and digital library projects can assist with the effective exploitation of this intellectual labour. With digitisation, effective knowledge management practices can be adopted, for example. By this means, value can be created and extracted from intellectual labour and profits can then be derived from this value. All this ensures the continued success of capitalism, whilst labour continues to be exploited, alienated and objectified, because value can only be created from labour. As Marx indicated: ‘the labourer, by virtue of his labour being of a specialised kind ... by the mere act of working, creates each instant an additional or new value’ (Marx, 1887, p. 201).

Thus, when we see digital projects gathering pace rapidly, we need to pause, and think and question. What is being gained from any one such project? Is digitisation always worthwhile? Does it necessarily mean that our information resources are being preserved and utilised more effectively? To what extent are commercial digital projects gathering pace and how is this likely to develop in the future (given that profit is the driving factor)? Could digitisation mean that few, if any, hard copies are available in the future? Could it mean that increasing amounts of material will not be available in the future, as technology changes, and the material is not transferred to the new medium? Might it lessen the value of scholarly publications (see, for example, Kalay’s chapter in this book). What are the likely health risks of these digitisation programmes? And who is likely to benefit from this digitisation process? As I outlined earlier in this chapter digitisation and digital libraries obviously can and do bring enormous benefits to society, particularly in regard to making information easily and readily available. However, questioning is also necessary, I think.

CONCLUSION

In conclusion, many different issues in regard to digital libraries have been discussed in this chapter. This includes a definition of digital libraries; the advantages of digitisation; the costs of digitisation; the UK Joint Information Systems Committee (JISC), and in particular, its work with digital preservation; traditional, digital, hybrid and virtual libraries; the library community and digitisation; e-books; a selection of different digital library projects on a global basis; gender issues, information technology and digitisation; and the social and political implications of digitisation and digital libraries.

Digitisation and digital libraries offer many new and exciting ways for us to be able to access material, but we also need to exercise some caution, and not just jump on every digital bandwagon.

NOTES

1 http://diligentproject.org/content/view/73/100/
DIGITAL LIBRARIES AND DIGITISATION

REFERENCES


Ruth Rikowski
Chandos Publishing, Oxford