EVALUATING AND MEASURING THE VALUE, USE AND IMPACT OF DIGITAL COLLECTIONS

EDITED BY LORNA M. HUGHES
Evaluating and Measuring the Value, Use and Impact of Digital Collections
Evaluating and Measuring the Value, Use and Impact of Digital Collections

Edited by
Lorna M. Hughes
Contents

Figures and tables ..............................................................................................................vii
Foreword  Christine L. Borgman ....................................................................................ix
Acknowledgements ........................................................................................................xii
The contributors ..............................................................................................................xiii

1 Introduction: the value, use and impact of digital collections ............................1
   Lorna M. Hughes

Part 1 Digital transformations in libraries, museums and archives 11

2 The digital library .......................................................................................................13
   Andrew Prescott

3 The digital museum ...................................................................................................35
   Claire Hudson

4 The digital archive .....................................................................................................49
   Gillian Oliver

Part 2 Understanding and measuring the use, impact and value of digital collections 61

5 A strategic approach to the understanding and evaluation of impact .......63
   Ben Showers

6 User needs in digitization .........................................................................................73
   Milena Dobrev, Andy O’Dwyer and Leo Konstantelos

7 Measuring impact and use: scholarly information-seeking behaviour ........85
   Claire Ross, Melissa Terras and Vera Motyckova

8 The value and impact of digitized resources for learning, teaching, research and enjoyment .................................................................103
   Simon Tanner

Part 3 Enhancing the future impact and value of digital collections 121

9 Using ICT methods and tools in arts and humanities research ......................123
   Lorna M. Hughes

10 Creating a research data infrastructure: policy and practicalities ................135
    Ann Borda and Lyle Winton

11 Improving sustainability of publicly funded digital resources ....................147
    David Robey

References and further reading ............................................................................157
Index ......................................................................................................................173
Figures and tables

Figures
Figure 3.1 Example of a PeoplePlayUK object story (copyright Victoria and Albert Museum) ..........................................................36
Figure 3.2 Theatre backcloth designed by Natalia Goncharova for a 1926 production of The Firebird at the Lyceum Theatre. The cloth measures over 10 metres by 15 metres and is consequently rarely displayed (copyright Victoria and Albert Museum) .................................43
Figure 3.3 Theatre backcloth designed by Natalia Goncharova on display in the V&A (copyright Victoria and Albert Museum) .........................43
Figure 3.4 Costume for Swan Lake, designed by Alexandre Golovine, Moscow 1901. A complex costume such as this requires thorough mounting and preparation in order to convey an accurate impression of its effect on stage (copyright Victoria and Albert Museum) ............................................................................47
Figure 4.1 At Island Command by Allan Barns-Graham. Reproduced by permission of Archives New Zealand (Archives New Zealand/Te Rua Mahara o te Kaawanatanga Wellington Office AAAC 898 NCWA 316) ..........................................................................................................................53
Figure 7.1 Respondents’ reasons for using the British Museum Collection Database Online .................................................................89
Figure 7.2 The range of roles in academia held by the respondents .................90
Figure 7.3 Categories of search in the British Museum Collection Database Online .........................................................................................91
Figure 7.4 Main search queries related to a specific object or group of objects ....91
Figure 7.5 Respondents’ visitation patterns to the Collection Database Online ...92
Figure 7.6 Potential uses of social media with the Collection Database Online ...97
Figure 8.1 How digital resources transform research ........................................105
Figure 8.2 The Old Bailey Online homepage (copyright Old Bailey Proceedings Online) ..................................................................................106
Figure 8.3 How digital projects have transformed scholarship ....................107
Figure 8.4 Jane Austen’s Fiction Manuscripts Digital Edition (copyright King’s College London) .................................................................................109
Figure 8.5 How digitized resources provide economic and educational benefits ....111
Figure 8.6 How digitized resources can optimize the research and teaching environment ..................................................................................113
Figure 8.7 The growth of digitized resources generates new and improved skills in students, scholars and the wider public.................................................................114
Figure 8.8 Digitized resources offer great support for personal participation in society..................................................................................................................116
Figure 8.9 The five modes of value for digitized resources.................................119
Figure 10.1 High-level architecture of typical virtual research environments and supporting infrastructure (copyright VERSI).........................................................142

Tables
Table 6.1 Types of user involvement in digitization............................................77
Table 6.2 User-related questions and methods .................................................79
Table 7.1 Respondents using COL by age distribution....................................90
Table 7.2 Respondents’ access points to the Collection Database Online........93
Table 7.3 Respondents’ opinions on design, layout and functionality of the Collection Database Online.................................................................94
Table 7.4 Responses to Collection Database Online potential areas for improvement..................................................................................................................94
Table 7.5 Respondents’ opinions on content in the Collection Database Online..95
Table 7.6 Respondents’ opinions on whether the Collection Database Online encourages a visit to the physical museum.......................................................96
Table 7.7 Key search terms provided for task-based searching scenario 1.........99
Table 7.8 Key search terms provided for task-based searching scenario 2........100
Table 10.1 Research data output types and data management planning areas.....140
Table 10.2 Retention strategies, costs and risks ..............................................143
Foreword

This set of chapters on digital collections fills an important gap in the professional literature of the memory institutions: libraries, museums, and archives. While much has been written on the evolution of digital scholarship, most analyses are written by and for scholars. These chapters are written by librarians, archivists and scholars engaged in building, assembling and digitizing content for a range of audiences, largely in the humanities. As noted in the acknowledgements, the book emerged from an expert seminar in e-research. The result is a coherent arrangement of chapters from a group of authors collaborating toward a common goal of identifying metrics for digital collections.

They survey developments, concerns, best practices and criteria for evaluation in a wide range of projects across the United Kingdom, Australia and New Zealand – where these authors currently are based – with reference to projects in the United States and elsewhere. The comparison of efforts in libraries, museums and archives, which in turn serve universities, schools, theatres and other environments, offers a rich set of case studies. A number of policy issues cut across these environments, most notably intellectual property rights and provenance.

Information professionals, managers and students alike will find much of value in this volume. The current environment of accountability is particularly problematic for the humanities. The ‘impact’ or value of collections may not be evident for years, decades, or even centuries after the origin of the materials, yet funders demand immediate economic indicators. These chapters take a balanced approach, acknowledging the trade-off in short and long term assessment of value, and to whom.

In all fields, the availability of scholarly content in digital form makes possible new research questions, methods and uses. The humanities especially have benefited from the ability to digitize historic documents, to mine large corpuses of texts, audio and images, and to assemble widely dispersed cultural objects into common repositories for comparison and analysis. Yet digitization is a means for scholarship and learning, not an end in itself. Careful assessments are required of trade-offs between usability and honesty to original form, between image quality, speed of access and cost, and between assorted other project-specific factors. Managers must identify their goals clearly to steer their way through the sea of standards that apply to digital projects. The orientation towards assessing use and users is particularly valuable, as it is often under-appreciated in digital projects.

Some of the projects represented here go well beyond scholarship, making
historical resources useful not only for research, teaching and learning, but also for ‘enjoyment’. Would that all digital projects keep the joy of discovery at the centre of attention!

Christine L. Borgman
Professor and Presidential Chair in Information Studies,
University of California, Los Angeles
Acknowledgements

Many of the papers in this volume were initially developed at an Expert Seminar hosted by the Centre for e-Research at King’s College London on 12 July 2010. Participants in that event presented their initial ideas for this book, and the discussions that followed shaped the content that has emerged. The participants were Sheila Anderson and Simon Tanner (King’s College London), Ann Borda and Lyle Winton (University of Melbourne), Jean-Claude Guédon (University of Montreal), Andrew Prescott (University of Glasgow), Claire Hudson (the Victoria and Albert Museum), Ben Showers (Joint Information Systems Committee), Milena Dobreva (University of Strathclyde), Claire Warwick (University College London) and Monica Bulger (University of Oxford). From the discussions at that event, ideas emerged about other strands that could be included in the book, and I am grateful to Melissa Terras, Claire Ross, Vera Motyckova, David Robey, Andy O’Dwyer, Leo Konstantelos and Gillian Oliver for contributing chapters and additional material that have shaped and completed the content in this volume. I would also like to thank Marilyn Deegan, Harold Short, Andrew Green, Arwel Jones and Avril Jones for their assistance and much-valued input to the project.

I am grateful to my former colleagues at King’s College, both in the Centre for e-Research and the Centre for Computing in the Humanities, for their collaboration over the five years I worked at King’s. I would also like to express my thanks to my new colleagues at the National Library of Wales, who have been enormously supportive and encouraging of this project since I took up my new post at the Library in 2011. The National Library of Wales has been a pioneer in digitization, and working among their digital collections, and the experts who create, manage and sustain them, has been a supportive and intellectually stimulating environment in which to complete the editing of this book.

Thanks are of course also due to all those who made images available for use in this volume.

I’d also like to thank the team at Facet, Sarah Busby and Lin Franklin, for being extremely helpful and encouraging at all stages of the development of the text.

Above all, I’d like to thank Russell, Archie and Lachlan for their patience and support.
The contributors

**Ann Borda** has held strategic roles in academic and research organizations for 15 years. Ann is currently Executive Director of the Australian state-funded Victorian eResearch Strategic Initiative (VeRSI), which is a multimillion dollar programme providing a coordinated approach to accelerating the uptake of e-research by the use of advanced information and communication technology (ICT) to enhance research outcomes, support collaboration and improve access to data and resources. Previously, Ann worked with the Joint Information Systems Committee (JISC) where she was responsible for the quality delivery of government-funded projects in building a UK wide e-infrastructure to support research. Ann is an Honorary Principal Fellow at the University of Melbourne and she has published in the areas of virtual research collaboration, data modelling and information system design.

**Christine L. Borgman** is Professor and Presidential Chair in Information Studies at the University of California, Los Angeles (UCLA). She is the author of more than 200 publications in the fields of information studies, computer science, and communication. Both of her sole-authored monographs, *Scholarship in the Digital Age: information, infrastructure, and the internet* (MIT Press, 2007) and *From Gutenberg to the Global Information Infrastructure: access to information in a networked world* (MIT Press, 2000), have won the Best Information Science Book of the Year award from the American Society for Information Science and Technology. She is a lead investigator for the Center for Embedded Networked Systems (CENS), a US National Science Foundation Science and Technology Center, where she conducts data practices research. Current collaborations include Monitoring, Modelling, and Memory and the Data Conservancy project. She is a member of the US National Academies’ Board on Research Data and Information, the US National CODATA and the Board of Directors of the Electronic Privacy Information Center, and a Fellow of the American Association for the Advancement of Science. She is also the 2011 recipient of the Paul Evan Peters Award from the Coalition for Networked Information, the Association for Research Libraries and EDUCAUSE, and the 2011 recipient of the Research in Information Science Award from the American Association of Information Science and Technology.

**Milena Dobrev**a is a research fellow at the School for Creative Technologies, University of Portsmouth. Before joining the School, she worked for the University of Strathclyde where she managed the DiSCmap project (JISC, 2010b) funded by JISC, studying priorities in digitization for UK higher education institutions from the point of view of
EVALUATING & MEASURING THE VALUE, USE & IMPACT OF DIGITAL COLLECTIONS

users (2008–9), the international Europeana user and functionality study (2009–10) and the local contribution to the Sustaining Heritage Access through Multivalent ArchiviNg (SHAMAN) project funded by FP7 of the European Community (EC) (2008–11). From 1990 she worked at IMI-BAS where she earned her PhD degree in informatics and applied mathematics. Milena was the founding head of the first Digitization Centre in Bulgaria (2004) and of the Humanities Informatics Department at IMI-BAS (2007). During 2002–5 she collaborated with the Humanities Advanced Technology and Information Institute (HATII) at the University of Glasgow on three Technology Watch reports on the application of new/core/emerging technologies in the cultural and scientific heritage sector for the DigiCULT project funded by the EC, and in 2007–8 she contributed to the work on the policy and quality domains within the DELOS Network of Excellence on Digital Libraries’ Digital Libraries Reference Model (DLRM). She also served as chair of the Bulgarian national committee of the Memory of the World programme of UNESCO. Her interests are in user studies for digital libraries, long-term preservation and e-culture.

Claire Hudson is Head of Collections Management for the Victoria and Albert (V&A) Museum’s Department of Theatre and Performance. She is a qualified librarian whose career began in public libraries; she then moved to the academic sector before she joined the V&A Theatre Museum (now the Department of Theatre and Performance) in 1987. Since then, her responsibilities have expanded into managing in their entirety the museum’s highly diverse performing arts collections. They include library material, archives, manuscripts, works of art, costume, museum objects, photographs and video collections, and form the UK’s national collection for the performing arts. With an emphasis on making the collections accessible to the research community and to the general public, her role also encompasses responsibility for digital projects. The department was an early adopter of digitization, beginning with a successful lottery-funded project in the late 1990s.

Claire is actively involved in national and international networks of performing arts curators, librarians and archivists, serving as President of SIBMAS – the International Association for Libraries and Museums of the Performing Arts – between 2004 and 2010.

Lorna M. Hughes is the University of Wales Professor of Digital Collections at the National Library of Wales. She took up her appointment in January 2011, and leads a research programme based around the Library’s digital collections. Prior to moving to the National Library of Wales in Aberystwyth, Lorna worked at King’s College, London, most recently as Deputy Director of the Centre for e-Research. From 2005 to 2008, she was Programme Manager of the Arts and Humanities Research Council (AHRC) ICT Methods Network, based in the Centre for Computing and the Humanities (CCH). She has worked in digital humanities at New York University, Arizona State University, Oxford University and Glasgow University. She is the author of *Digitizing Collections: strategic issues for the information manager* (Hughes, 2004) and the co-editor of *Virtual Representations of the...*
THE CONTRIBUTORS

Past (Greengrass and Hughes, 2008). She is the Secretary of the Alliance of Digital Humanities Organizations (ADHO).

Leo Konstantelos is a Preservation Researcher in HATII at the University of Glasgow and in the School of Creative Technologies at the University of Portsmouth. He holds a PhD on user requirements for born-digital art in digital library environments, and an MSc in information technology. Leo is the HATII Principal Investigator of the JISC-funded Preservation Of Complex Objects Symposia (POCOS) project. He has worked extensively in digital preservation projects, delivering innovative research in the area of digital/new media arts preservation requirements and in alternative manners for preserving dynamic and interactive material. He co-authored the Preservation and Long-term Access through NETworked Services (Planets) report on emerging digital art characterisation techniques (McHugh, Konstantelos and Barr, 2009), producing a vocabulary for context classification of dynamic, interactive and ephemeral digital artworks. He has worked for the EC-funded project SHAMAN. Leo recently joined the Keeping Emulation Environments Portable (KEEP) project team at the University of Portsmouth.

Vera Motyckova completed an MA in electronic communication and publishing at University College London (UCL) in 2010. During her studies she carried out research in collaboration with UCL and the British Museum focusing on the non-academic users of the British Museum’s Collection Database Online. She also worked at the British Museum on the development of new online research catalogues. Currently, Vera is part of the online editorial team at the British Broadcasting Corporation (BBC) in the Learning Department.

Andrew O’Dwyer is a Technologist/Project Manager at the BBC, working on digitization activities for preservation and access to audiovisual collections. He also works on a number of European Union (EU) collaborative projects to bring archives online for public and academic use. He is a member of the Television Studies Commission of FIAT, www.fiatffra.org, promoting academic use of audiovisual material, and a member of the European Television History Network (http://cms.hum.uu.nl/ethn). He is also a contributing author of the book A European Television History (Bignell and Fickers, 2008).

Gillian Oliver is currently Senior Lecturer in Archives and Records Management at the School of Information Management, Victoria University of Wellington, New Zealand. Her doctoral degree is from Monash University, Melbourne, Australia. Her professional practice background spans libraries (mainly in the health care sector), records management and archives. Most recently she was actively involved in establishing digital archiving capability at Archives New Zealand. Her teaching experience includes the development and delivery of online education for records and information management at The Open Polytechnic of New Zealand. Her research interests encompass record-keeping perspectives on digital preservation and organizational cultures.
EVALUATING & MEASURING THE VALUE, USE & IMPACT OF DIGITAL COLLECTIONS

Andrew Prescott is Professor of Digital Humanities at King’s College London, and until recently was Director of Research at the Humanities Advanced Technology and Information Institute, University of Glasgow. He was formerly a curator in the Department of Manuscripts at the British Library, where he was a library contact for a number of pioneering digitization projects, such as Electronic Beowulf. He edited, with Leona Carpenter and Simon Shaw, the account of the British Library’s first digital and network programme, *Towards the Digital Library: the British Library’s Initiatives for Access programme* (Carpenter, Prescott and Shaw, 1998). From 2000 to 2007, he worked in the Humanities Research Institute, University of Sheffield, where he was Founding Director of the Centre for Research into Freemasonry, the first academic centre in the UK devoted to the history of Freemasonry. From 2007 to 2010, Andrew was Librarian of the University of Wales Lampeter, serving latterly as Pro Vice Chancellor. His publications include *English Historical Documents* (Prescott, 1988), *The British Inheritance* (Hallam and Prescott, 1999) and *The Benedictional of St Aethelwold* (Prescott, 2002).

David Robey currently holds the position of Arts and Humanities Consultant at the Oxford e-Research Centre. Until recently he was Director of the Arts and Humanities Research Council’s ICT in Arts and Humanities Research Programme. He has been University Lecturer in Italian at Oxford and Fellow of Wolfson College, then Professor of Italian at Manchester, then Professor of Italian at Reading. He has published on 15th-century Italian humanism (educational and poetic theory), language and style in Dante and Renaissance narrative poetry, the computer analysis of literature, and modern critical theory. He has authored a computer-based study on sound and structure in Dante’s *Divine Comedy* (Robey, 2000), later extending this work to include the major narrative poems of the Italian Renaissance. He has also completed a major online analytical database on ‘Sound and Metre in Italian Narrative Verse’ ([www.italianverse.rdg.ac.uk](http://www.italianverse.rdg.ac.uk)). He was joint editor of *The Oxford Companion to Italian Literature* (Hainsworth and Robey, 2002, now translated as the *Enciclopedia della Letteratura Italiana*, Oxford/Zanichelli), and joint author of the forthcoming *Very Short Introduction to Italian Literature* (Hainsworth and Robey).

Claire Ross is a PhD student and researcher at the UCL Centre for Digital Humanities. Her research focuses on user evaluation and user centric design of social media applications, and digital museum content and online collections. Formerly, she held the position of E-Learning Development Project Manager working on a collaborative project with the University of Exeter and Geevor Tin Mine Museum. She is Chair of the Digital Learning Network, formerly the E-Learning Group For Museums, Libraries and Archives. She is also a member of the committee for Ignite London. Her research interests include usability studies, web 2.0 applications, social media, museum e-learning, digital heritage and digital repositories.

Ben Showers is a programme manager with the Digital Infrastructure team at JISC with responsibility for the Library Infrastructure programme. Ben manages a number of...
programmes of work, including library systems, user experience, mobile library services and shared services. Prior to this Ben worked with the e-Content team at JISC, working on a number of digitization and digital scholarly resource programmes helping to explore innovative new models of content creation and curation including crowdsourcing and user-generated content. Ben holds an MA in philosophy and an MA in library and information science from University College, London.

Simon Tanner is the Director of Digital Consultancy at King’s College London (www.kdes.kcl.ac.uk). He is also a senior academic in the Department for Digital Humanities at King’s and co-Director for the MA in digital asset management. He authored the book Digital Futures: strategies for the information age with Marilyn Deegan (Deegan and Tanner, 2001). Simon previously led in information technology (IT), management and library roles for the Higher Education Digitisation Service (Senior Consultant), Loughborough University (Systems Manager) and Rolls-Royce and Associates (Head of Library Services).

Melissa Terras is the Reader in Electronic Communication in the Department of Information Studies, University College London. She has a background in classical art history and English literature, and computing science. Her doctorate (University of Oxford) examined how to use advanced information engineering technologies to interpret and read the Vindolanda texts. Publications include Image to Interpretation: intelligent systems to aid historians in the reading of the Vindolanda texts (Terras, 2006) and Digital Images for the Information Professional (Terras, 2008). She is a general editor of DHQ: Digital Humanities Quarterly and the Secretary of the Association of Literary and Linguistic Computing. Her research focuses on the use of computational techniques to enable research in the arts and humanities that would otherwise be impossible.

Lyle Winton is eResearch Analyst and Policy Manager at VeRSI, University of Melbourne. His responsibilities include managing engagements with VeRSI stakeholders and the research community, exploring opportunities and solutions that align with VeRSI aims (advancing research, collaboration and sharing through ICT and e-research practice). His general interests lie in research informatics, research data management, distributed collaboration, service oriented architecture, and training and education with respect to all of these areas.
1

Introduction: the value, use and impact of digital collections

Lorna M. Hughes

Background and context
The aims of this volume
A key motivation for developing this volume was the need to address the ‘use’, ‘value’ and ‘impact’ of digital collections in the context of an expanding mass of digital content with tremendous potential. Specifically:

- How can we understand how digital collections are being used, and by whom?
- How do we assess their value, and add value over time, in order to make decisions about which collections to digitize or make available, and how?
- How do we assess their impact on scholarship, on knowledge transfer and on information management and access?
- How do we ‘measure’ value? What can be measured, and how?
- Is it possible to ensure their sustainability, value and impact over time?
- How might we apply indicators of use, value and impact to inform funding decisions and policy making for the future?

The explosion of digital initiatives
This volume of essays is, in many respects, a follow up and companion to Digitizing Collections: strategic issues for the information manager (Hughes, 2004). Since its publication, digitization initiatives have continued at a tremendous pace in libraries, archives and museums, as well as in higher education. Digitization of existing library, museum and archive collections is still a major priority, where funding can still be found for these initiatives. The National Library of Wales is continuing to digitize two million pages of historic Welsh newspapers and journals in a three-year project that will conclude in 2012, with funding from the Welsh Government. The British Library is partnering online publisher Brightsolid to digitize up to 40 million pages of newspapers. In the USA, the Smithsonian Institute has a remit to digitize its entire collection: a challenge considering that the Smithsonian is home to 137 million objects, 100,000 cubic feet of archival material and 1.8 million library volumes. In the UK, the Joint Information Systems Committee (JISC) has launched a new phase of digitization via its e-Content programme, for projects that will be completed in 2012. This is the era of Google, and mass digitization initiatives to put our cultural heritage online are flourishing. This has been described as a ‘data deluge’ (Hey and Trefethen, 2005), and it has a huge impact on scholarship, teaching and public engagement.
Digitization in an economic downturn
This is an auspicious time to take stock of this mass of digital content, and consider its impact, value and use. The global economic decline that began in 2007 has led to serious cuts in funding for almost all humanities and cultural heritage initiatives, including the development of, and support for, digital collections. In the arts, humanities and cultural heritage world, especially in the UK, the threats to funding have become a reality, with the closure of the Museums, Libraries and Archives Council (MLA), cuts to the Arts Council, and to Research Council funding overall. In the USA, there have been calls to withdraw federal funding from the National Endowment for the Humanities. This economic ‘austerity’ has created significant institutional and societal pressures on cultural heritage and higher education organizations.

Partly as a consequence of the reduction in funding, we have seen a sharper emphasis on the need to demonstrate the ‘impact’ of publicly funded resources and research, as a means of quantifying the value of the investment in their creation. Research councils and funding agencies, notably the Arts and Humanities Research Council (AHRC) in the UK, have placed an increasing emphasis on ‘impact’ and ‘evidence of value’ of all research that they fund (AHRC, 2006) for several years. This focus on understanding the ‘evidence of value and impact’ of digital research and collections in the arts and humanities was one of the reasons that AHRC ICT Methods Network was funded from 2005 to 2008. This was a national organization that provided a forum for the exchange and dissemination of expertise in the use of information and communication technologies (ICTs) for arts and humanities research. In its final evaluation report, the Methods Network was able to provide a considerable body of evidence that there was indeed ‘evidence of value’ of the use of digital collections for scholarship and research in the arts and humanities:

The new research that has been enabled by ICT ... has depended upon the development of new kinds of resources, such as large corpora in literary, linguistic, musicological, and television and film studies domains, the digitization and digital-encoded representation of materials in classics, history, literature and history of art, and the creation of databases in archaeology and the performing arts. This recognition that the future generations of scholarship in the arts and humanities will depend upon the accessibility of a vast array of digital resources in digital form is becoming more widespread. (Hockey and Ross, 2008)

Since then, the Browne review (Browne, 2010) has called for evidence of the ‘value’ of the arts and humanities to society, and there is increasing pressure for scholarly research to demonstrate economic and social impact, despite the fact that the ‘economic benefit of the arts and humanities’ is a topic for which there is little hard evidence at this time. Nonetheless, the AHRC has produced two publications that set out the arguments for the ‘value’ of the arts and humanities, in the recent publication edited by Jonathan Bate, The Public Value of the Humanities (Bate, 2010) and in the AHRC report Leading the World: the economic impact of UK arts and humanities research (AHRC, 2009). A noteworthy web-based initiative, 4Humanities (http://humanistica.ualberta.ca), organized by leading digital humanities scholars, has also been set up as an advocacy organization for the value of digital collections and methods in the humanities.
The need for open and useable digital collections

Another recent development that provides part of the background to this volume is the increasing focus on the value and use of ‘open’ scholarly resources for research. While publicly funded research outputs are intended to be freely available (even if this is not always the case, and as Robey discusses in this volume, the demise of the Arts and Humanities Data Service in the UK in 2008 has made it harder for publicly funded resources to be used and sustained), digitization that has been funded by commercial entities is frequently subject to licensing and use restrictions. Notably, in the UK the digitization and redistribution of the census materials and similar ‘name rich’ resources has been made possible by commercial entities including Ancestry.com and Find My Past.

While these efforts have digitized and made available an enormous range of primary source materials, there is some concern that resources that are not ‘open’ are less valuable for scholarship. This is of particular concern, as the availability of large-scale, distributed collections allows new approaches to scholarship in a number of disciplinary areas. The idea of taking a ‘big picture’ approach to historical and cultural issues, working with large-scale data across disciplines, is gaining ground in scholarly enquiry once more. This is the type of research highlighted in the American Council of Learned Societies report, emphasizing the opportunities ‘to reintegrate the cultural record, connecting its disparate parts and making the resulting whole available to one and all, over the network’ (ACLS, 2006). Underpinning this re-integration, of course, is the principle of freely available and open data for aggregation, use and reuse: ‘the full range of online content needs to be made available to all, quickly, easily and in a form appropriate to individuals needs’ (JISC Strategic Content Alliance, 2010).

Some organizations have made the principle of providing freely available digital content where possible (see Prescott’s discussion of the National Library of Wales in this volume), but this is not always possible, especially in the current economic climate. There are various approaches: can data be chargeable for a time to recover the costs in making it? Under what conditions can cultural heritage content be monetized? This volume is not about business models for digital content, but there is an implicit understanding that ‘price’ and ‘value’ could become interconnected – people may have to pay for valuable resources. There are also issues related to sustainability – the digital content must be sustained as long as charging models are in place – which adds to their cost. In a related development, commercial entities have raised concerns about digitization of out of copyright materials in libraries (Sabbagh, 2010). However, these discussions lack the evidence of the actual economic ‘value’ of digital resources, and work is under way to address this balance, developing a ‘deeper understanding of the social and economic impact of digitization’ (Hargreaves, 2011).

The cost of digital creativity

While the economic data is incomplete – and this volume is not a quantitative analysis of business models for digitization – one thing that is very clear is that digital collections are expensive to develop, manage and sustain over the long term. This has been
documented in two reports commissioned by the AHRC: its Resource Enhancement Scheme (AHRC, 2005) and Sustainability of Digital Outputs from AHRC Resource Enhancement Projects (Denbo, Haskins and Robey, 2008). The latter report demonstrated that ‘The average grant to projects with digital outputs was £228,155, as against £153,090 to those without: the total amount invested in those with digital outputs was therefore almost £40m.’ Furthermore, there are hidden costs to developing digital resources – institutional overheads in keeping collections online and visible, and in the allocation of resources to digital programmes, frequently at the cost of other activities. In academia, the costs are less hidden – there is a body of evidence (see, for example, Raban, 2007) demonstrating that the development of digital resources by academics can hinder their career progress, and even jeopardize promotion or tenure. It is notable the research findings presented in this volume by Ross, Terras and Motyckova show that the largest numbers of academic users of the British Museum Collections online database were either postgraduate students or professors – either those just starting out, or those secure in their promotions. Impact, of course, can be negative as well as positive, so it is timely to assess the pressures that developing, sustaining and using digital collections can place on organizations and individuals.

Protecting the digital investment
Given the current economic climate, and the pressure on funding for digital collections, addressing these issues is crucial in order to make the case for protecting existing digital heritage, and for increased digitization. New digital projects must have compelling and visible impact. It is necessary to be creative about the sort of funding available, and to look further afield, both internationally and in disciplinary focus, for funding. In this regard, the economic downturn has been a way of incentivizing collaboration, enhancing existing resources, and aggregating content, as well as consolidating expertise in a few centres. Research Council funding in the UK has also been mostly contingent on developing digital resources that are driven by research challenges, rather than resource enhancement or access, which may well create resources that have more demonstrable ‘value’ for scholarship. This has been a ‘carrot and stick’ approach – the carrot being the opportunity to apply for the limited funding still available, the stick being that this is the only way that funding can be awarded. Equally pragmatically, as it becomes more difficult to get funding for digital collections, there is a need to make better use of those we have, and to ensure that they are fully embedded in research, teaching and public engagement. There is a need to show clearly – to funders, to the public and to those responsible for their long-term sustainability – that digital resources can enable scholarship that generates new research questions and findings; that they make it easier and more efficient to carry out ‘traditional’ scholarship through better and enhanced access to resources (Lehmann and Renfro (1991) and Wiberley (2000) suggest that humanities scholars are receptive to technology as long as it demonstrates adequate savings in time or effort); and that they extend the evidence base for research. By expanding the cultural heritage collections available to the public, they also have economic and societal impacts. However, this
evidence is not readily available, as outlined in the keynote address at the Digital Humanities Conference in 2010 (Terras, 2010), where the case was forcefully made for better, more readily available exemplars of the use and value of digital collections and research.

Evidence of use and impact
Some evidence exists for the use of digital resources for research, education and public engagement, thanks to initiatives like the JISC e-Content programme; the Methods Network; and the National Endowment for the Humanities Office of Digital Humanities, notably the ‘Digging into Data’ initiative, which seeks to use large digital corpora to address key research challenges in the humanities (see hwww.diggingintodata.org). The AHRC ICT Methods Network produced a body of evidence (documented in the final evaluation of the programme) that demonstrated that digital resources were having a transformative effect on scholarship, and that this was a vibrant, fast-expanding area. However, as the final evaluation stressed, the full impact of the Methods Network and, indeed, the impact of digital resources in the arts and humanities, will be over the longer term (Hockey and Ross, 2008). But, since the demise of the Network, there has been no attempt to systematically address these issues across the disciplines. A longitudinal, multi-disciplinary study on use and impact of digital resources is still needed.

Value
One of the key issues is being able to provide a definition of ‘value’. Digital resources are valuable to different audiences for different reasons, and some value may not be realized immediately. Digital collections come about for different reasons. Many research projects, for example, have produced digital images or digital text as a by-product of scholarship, and a need to put these digital images online for public access may not be the first priority of the project team. For example, ‘The Visual Culture of Wales’ Project (based at the University of Wales Centre for Advanced Welsh and Celtic Studies) was funded mainly through the Arts Council of Wales Lottery Unit and the University of Wales from 1996 to 2003, as a publication project. The photographic archive contains digitized photography of around 3000 works of art, dating from the fifth century up to the 1960s, in a wide variety of media. The database for the images was subsequently made available online, with minimal funding, as a by-product of the research publication project. Similarly, name-rich sources (parish records, census records, baptism records) have been digitized by the Church of Latter Day Saints for theological reasons. The digital files may be invaluable for family history, or demographic research, but this was certainly not the motivation for their creation. Other digital resources, including born-digital materials, are being preserved for future use – the Library of Congress, for example, has archived the Twitter archive, which could prove just as useful to scholars of the early 21st century as 19th-century newspapers online have proved to political, social and cultural historians of the newspaper era. In this regard, digital resources are just like any other cultural artefact – historical resources that are used for research were seldom
EVALUATING & MEASURING THE VALUE, USE & IMPACT OF DIGITAL COLLECTIONS

intended to be ‘historical resources’; they were administrative records, ephemera or cultural artefacts. They tell us more than they were ever intended to. This is a useful truth to apply to digital content. As we create more and more digital content – as part of conscious attempts to digitize analogue materials, as by-products of research or as part of our daily digital mode for work and communication – we are creating, maintaining and sustaining ‘digital content’ that may have value to future scholars.

A digital text may be valuable to a scholar because it enables the use of text mining tools to undertake historic research, or it may be valuable to family historians as it mentions an ancestor, or to scholars of material culture through its description of objects. It may gain value if linked to other digital content through ‘virtual reunification’, where collections held in disparate archives around the world can be combined in digital facsimiles. Value is subjective, changes over time and has different meanings that are contingent on external factors. Value of digital collections, and digital humanities in general, is particularly difficult to assess as ‘Digital Humanities is not a unified field but an array of convergent practices’ in multimedia configurations; using digital tools, techniques and media have altered the production and dissemination of knowledge in the arts, human and social sciences, creating ‘digital models of scholarly discourse for the newly emergent public spheres of the present era (the WWW, the blogosphere, digital libraries, etc.), to model excellence and innovation in these domains, and to facilitate the formation of networks of knowledge production, exchange, and dissemination that are, at once, global and local’ (Schnapp and Presner, 2009).

This is, of course, a common issue across the arts and humanities: funded research often does not show its ‘value’ until it is broadly disseminated, shared, cited and becomes common across the disciplines. This can take time. It is the same with digital resources, and funders often fail to appreciate that the ‘value’ of digital collections and the scholarship they enable may take time to emerge. The chapters by Prescott, Hudson and Oliver are interesting in this regard as they take a longer view of the impact of digital collections and programmes on the institutions that host them.

Some of the factors that make digital collections ‘valuable’ to libraries, archives, museums and higher education were addressed in Digitizing Collections: strategic issues for the information manager (Hughes, 2004). These included:

- access, both broader access to a global audience via the internet, and enhanced access, by making aspects of collections searchable, findable and linked to related materials
- supporting preservation, by providing digital surrogates of rare and fragile materials
- collections development, by enabling organizations to develop cataloguing and records management around digital objects, and by enabling the ‘virtual reunification’ of collections that are physically separated
- institutional and strategic benefits, such as professional development of staff, the prestige and PR value to the institution; and enabling the institution to fulfil its goals of access and outreach
Hughes: Introduction

- supporting research and education across the disciplines.

The book also investigated the impact of digital collections on institutions, including:

- the need for new business models to support and develop digital collections, and how this was forcing institutions to consider the institutional costs and benefits of digital collections, grappling with the contradiction that there can be indirect cost savings from digital delivery of services, but that these are offset by the increased costs of digital access and preservation.
- the resources and intellectual implications of changes to the way that information is used and managed.
- the implications of supporting entirely new approaches to scholarship and access.

The book concluded that there was no definitive evidence base that could provide concrete numbers about the economic ‘value’ of digital collections, but there were some interesting debates around the topics at the time, for example, ‘The Economics of Digitizing Library and Other Cultural Materials’, which identified the ‘costs’ of digitization: institutional, technological and legal (Waters, 2004).

The structure and content of this volume

This volume has three parts. The first, ‘Digital transformations in libraries, museums and archives’, describes the use and impact of digital collections in libraries, museums and archives. Andrew Prescott, Claire Hudson and Gillian Oliver provide valuable examples that capture practice and experience in a variety of organizations, discussing ways that the digital revolution has transformed the mission and organization of services and collections. The second part, ‘Understanding and measuring the use, impact and value of digital collections’, presents different approaches to measuring and understanding value. Ben Showers describes some approaches developed through the JISC e-Content programme to measure impact and to embed digital collections more thoughtfully within user communities. Milena Dobreva, Andy O'Dwyer and Leo Konstantelos present a series of approaches to user testing that can inform the development of more valuable resources, and discuss how this can impact the development of business models. Claire Ross, Melissa Terras and Vera Motyckova discuss a survey of users of the British Museum Collections Online Database, and develop an analysis of the use of digital content by academic researchers. Simon Tanner concludes this section with an overview of the opportunities and impacts that digitized resources have made for learning, teaching, research and society.

The third part, ‘Enhancing the future impact and value of digital collections’, presents some approaches that can add value to digital collections: Lorna M. Hughes describes ways in which they can effect transformative research via the use of ICT tools and methods; Ann Borda and Lyle Winton discuss the ways in which e-infrastructures can enable collaborative scholarship and shared resources, considering how preservation and
EVALUATING & MEASURING THE VALUE, USE & IMPACT OF DIGITAL COLLECTIONS

research infrastructures and strategies developed for the sciences can be applicable to other disciplines, and citing the exemplary work that has been funded in Australia in order to support this; and David Robey discusses sustainability and the long-term value of digital resources.

A concept with considerable impact is the ‘holistic’ approach to digitization, which sees the whole digital life cycle as significant and interdependent, rather than as a series of individual phases or stages. For illustrations of the concept, see, for example, the work of the Digital Curation Coalition in the UK (www.dcc.ac.uk), and the recent report by the Strategic Content Alliance, Developing Digipedia: a guide to the digital content lifecycle (JISC Strategic Content Alliance, 2010). Taking this approach at the outset of a digitization project should make the resource more ‘valuable’, as it will enable use and reuse to be embedded into the resource at each phase of development (Hughes, 2008). However, many digitization projects do not develop this approach. Therefore, at a practical level, this volume covers a number of useful and valuable things that the managers of digital collections can do at any stage of the digital life cycle in order to develop a better understanding of, and to increase, the use, value and impact of their resource. In terms of documenting resources, making their impact more visible and providing badly needed use cases (and guidelines for use, which are equally important) to demonstrate the transformative effect of digital collections, the five modes of value suggested by Tanner are an extremely useful taxonomy that could be useful for categorizing and documenting projects and their outputs. This model could well be adopted as a means of documenting ‘impact’ while projects are in progress, or applying to published digital collections retrospectively. Similarly, the ‘impact framework’ that Showers describes could well be adopted by project managers as an internal documentation exercise, or for dissemination. Showers also describes the implementation of the useful Toolkit for the Impact of Digitised Scholarly Resources (TIDSR) toolkit by projects funded by the JISC e-Content Impact and Embedding Programme. This is an easily adoptable methodology for using both qualitative and quantitative methods for assessing the use of a digital resource, which can produce useful findings. Similarly, the information manager might adopt the approach described in detail by Ross, Terras and Motyckova to understand user information-seeking behaviours, in order to get a deeper sense of usage patterns. Ross, Terras and Motyckova, and Showers describe approaches that can be applied at any time after a resource has been published, while Dobrev, O’Dwyer and Konstantelos describe in detail the value of applying usability and user engagement methods into the development of a digital resource. The importance of this cannot be underestimated, as usability is inconsistently factored into the project development cycle.

However, it is also possible to make resources more valuable and useful after they have been publicly available for some time. The JISC projects described by Showers are examples of ways in which projects can be enhanced by becoming more embedded in research, teaching and public engagement. Hughes describes some of the ways in which digital resources can underpin scholarly use through the application of ICT tools and methods. From a strategic perspective, Bordura describes how data management and use
can be supported through research infrastructures and in virtual research environments, to foster collaborative and individual use and reuse of existing collections. At the heart of all this work is of course the expectation that resources are open to allow use and reuse for unforeseen purposes. It also supposes that the resources will be available to all. The FEDORA (Flexible Extensible Digital Object and Repository Architecture) commons community (http://fedora-commons.org) now favours the term ‘durability’ to describe robust digital resources that are sustained and preserved over the long term for use by many communities and purposes – and this approach would indeed add ‘value’ to digital resources. Sustaining durable digital resources over the long term will add institutional overhead costs to digital collections. However, the visibility of digital collections means that they must be addressed, unlike the forgotten ‘unseen’ costs to libraries of preserving and managing the monograph.

Conducting user assessments and studies, developing use cases and developing ways of embedding digital collections more effectively in teaching and research are, of course, activities that require the time and effort of staff, and therefore additional investment after a resource has been published. This is often a daunting prospect – digital projects are usually developed through short-term funding, with staff on short-term contracts. Therefore, after they are launched, they are seldom given any more than the most cursory technical attention. The projects described by Showers have attempted to redress this, by providing small amounts of JISC funding to ‘revisit’ projects to enhance usability or embedding, and this approach is an extremely useful one. Institutions may look to ways that they can themselves underwrite this sort of activity – possibly by using volunteers, interns or students to work on the enhancement of digital collections. It is also the sort of activity that users themselves can contribute to – every project has a core community of users, and identifying who they are and working with them to enhance resources could be a valuable opportunity to create a ‘virtuous circle’ of collaboration, engagement and outreach, integrating people and collections.

At the heart of the volume is the theme of digitization as a process with many stages – a digital life cycle – during which issues related to use, value and impact can be factored in. Digital culture is far more complex and encompasses more realities than textual cultures. At an Expert Seminar at King’s College London where preliminary ideas for the volume were discussed, Jean-Claude Guédon introduced this idea with the useful analogy that working with digital documents brings us back to the ways of working that were more familiar in a pre-print time, especially that of individual processing of documents, rather than mass printing. The production of medieval manuscripts is also a useful metaphor for the process of developing digital materials. In the world of manuscripts, communities took responsibility for maintaining texts carefully as a collective endeavour, as texts and the effort that had gone into their creation were easily lost. Texts were maintained due to the efforts of the community to sustain them. Churches had a role as supporters of these communities to preserve texts, and community involvement was essential for preservation: an early prototype of ‘collective intelligence’ before Wikipedia. ‘Use’ was key to sustainability of the resources – texts from antiquity were
preserved as they were used as grammars, demonstrating how a community often adds value to material. To continue the analogy, digital collections need to be tended, par the concept of digital collections as ‘processes’ rather than ‘products’. Diffuse communities use collections in different ways – and memory organizations create these relationships between the objects and the public.

Conclusion: future strategic directions
There are some key strategic themes discussed in this volume that still need to be addressed. In order to more fully embed use, value and impact into the development of future digital collections, strategic input from funders is required, especially with regard to recommending that projects they fund take a more robust approach to evaluating and demonstrating the impact, value and use of digital collections. Implementing a form of ‘peer review’ for digital resources, as suggested here by Robey, could contribute to this agenda. Documenting and measuring value and impact should be built into project planning and funding; an impact assessment should be carried out while the project is in development, and then funding made available for a summative piece several years after the publication of the digital resources. The value, impact and use of collections take time to evolve and to be understood, and this needs to be reconciled in a world of responsive, short-term funding opportunities. Similarly, while the underlying technical infrastructure should support aggregation and use of content across collections and organizations, a human infrastructure is also needed enabling and supporting the use, citation and durability of digital collections. This sort of collaboration, through networks and partnerships, must be supported over the long term to develop and disseminate the evidence of the use and value and impact of digital collections.

Notes
1 Part of the Digital Futures series, edited by Marilyn Deegan and Simon Tanner.
3 See www.smartplanet.com/people/blog/pure-genius/smithsonian-secretary-plans-to-make-massive-collection-available-to-all/3126/.
4 See the website of the National Humanities Alliance for updates on these and other issues: www.nh alliance.org/.
5 An AHRC funded network (http://historicweather.cerch.kcl.ac.uk/) is investigating integration of historical archives for the analysis of weather records and data over time, working in partnership with the Met Office ACRE project, which is carrying out digitization and analysis of historic weather data, primarily from ships’ logbooks (www.met-acre.org).
6 The JISC funded Shipping Archives and Integrated Logbooks of Ships (SAILS), a linked data project, was extremely successful in working with maritime history MA students to develop transcriptions of data for the project (http://sailsproject.cerch.kcl.ac.uk).
Part 1
Digital transformations in libraries, museums and archives
2

The digital library

Andrew Prescott

Introduction
This chapter takes the long view of the digital library, and reflects on the historical context of library developments, identifying some of the key ways in which digital transformations have enabled libraries around the world to educate, inform and delight their readers. These developments are then considered in a case study that highlights ways in which these changes have been implemented at the National Library of Wales, illustrating the ways in which these changes are impacting research in the arts and humanities.

A celebration of the library
The Argentinian bibliophile Alberto Manguel possesses a huge private library, which he has installed in a 15th-century barn at his home in the Loire Valley. In his book The Library at Night (2008), Manguel describes the magical atmosphere of his library at night-time, when pools of light over the library desks expunge the outside world, so that sounds become muffled and the very thoughts in the books seem clearer and louder. The comforting smells of the wooden shelves and the musky leather bindings permeate the library, and these smells seem to convey connections with ancient human knowledge and dreams. In this private world of public knowledge, connected through its books to countless other cultures and literatures, Manguel’s mind roams through other great libraries, from the great building at Alexandria to the private collections of authors such as Dickens or Borges.

Manguel’s book is a celebration of the library as a space for study, reflection, exploration, inspiration, privacy and sharing. Whether it is an ancient room lined with oak bookcases, a large and extravagant Edwardian municipal building, or a clean and functional modern block, the library is an evocative space where the mind can take wing, finding new connections and discerning new possibilities. It is a space that is at once private and public, where people of all social classes, intellectual accomplishments and enthusiasms can discover new thoughts and intellectual vistas. The library is one of the most potent of all spaces created by humanity, eliciting deeply personal reactions from all those who encounter it, from members of the public to wizened researchers. The architect of the British Library at St Pancras, Colin St John Wilson, expressed this library fever very well when he declared that ‘To every scholar the library is a realm of secret topography’ (Losh, 2004, 378). For Wilson, a library was the modern equivalent of a
medieval cathedral, whose sacral potency was enhanced rather than detracted by modern technological developments. He declared that ‘A great library is like a coral reef whose exquisite structure as it grows proliferates a living network of connectedness, and its ramification is all of a piece, like knowledge itself – the knowledge that bridges the endless curiosity of the human mind, from the first pictogram to the latest microchip’ (McCarthy, 2008).

The changing library
Libraries are an aspect of the ecosystem of humanities scholarship where the impact of digital and networking technologies has already been profound, and if there is one area in which the working practices of humanities scholars have been utterly transformed as a result of new technologies, it is the use of the library. Humanities academics who are, like me, in their late fifties have witnessed an astonishing transformation of the library space. At the heart of the traditional library space with which I grew up was the physical catalogue – as can be seen from the way in which the huge General Catalogue of the British Museum stood in a great arc at the centre of the Round Reading Room, like the cortex of a huge brain. The catalogues of large British libraries like the British Museum were published, but for many smaller academic libraries an up-to-date catalogue was only available on-site and the availability of particular items could only be established by a personal visit, which in the case of libraries at a distance might first necessitate lengthy correspondence. From the 1970s, the gradual conversion of catalogues into machine-readable form, then their availability online, increasingly through aggregated resources such as COPAC (the online public access catalogue of CURL, the Consortium of University Research Libraries) for British research libraries, or the international WorldCat developed by the OCLC (Online Computer Library Center) library consortium, transformed the logistics of tracing and consulting the rare books, manuscripts and archives with which many humanities scholars work. The libraries I used as a young scholar were dominated by serried dusty ranks of printed journals, and a journal search was a lengthy and physically arduous prerequisite of any new research project. Today, I rarely consult a printed journal, since articles are so much more readily and conveniently made available through a variety of repositories ranging from multi-publisher archives such as JSTOR and Project Muse, gateways developed by journal suppliers such as SwetsWise and the packages offered by the journal publishers themselves. Above all, as a historian I now have direct access to images and editions of primary sources online. If I wanted as a young man to consult newspapers, it usually necessitated a journey to the out-of-the-way suburban idyll of the British Library Newspaper Reading Room at Colindale. Nowadays, I can sit late at night, hunched at my laptop in a pool of light as evocative as that of Manguel’s library in the French countryside, ranging through newspapers covering every conceivable quirk and foible of British society in the 18th century.

Of course, it is not just access to British materials that have been transformed. One of the reasons that I became a British historian was because of the balance of library
provision when I was a young man; the barriers for a scholar researching European history in the London of the 1970s were considerable. These restrictions have been greatly eased. If I tire of British periodicals, I can turn to the French titles made available by the Bibliothèque Nationale de France through Gallica or cross the Atlantic to trace the footsteps of Wyatt Earp in the Tombstone Epitaph available through the Library of Congress. The Victorian founders of the modern British Library such as Anthony Panizzi and his lieutenant Thomas Watts dreamed of creating a great imperial library in London in which the best collection of English books in the world would be allied to the best collections of Russian books outside Russia, the best collection of German books outside Germany, and so on (Hill, 1953, 38). In building the circular British Museum Reading Room, surrounded by state-of-the-art industrial iron bookshelves designed to accommodate such a huge library and make it speedily available, Panizzi and Watts created a potent cultural symbol of Victorian Britain. This was predicated on the idea that at the heart of the largest city in the world should be a library that gathered together the world’s culture. This collection instinct – the idea that the function of a library was to pull in material from as far afield as possible – was fundamental to the traditional library. Although in the pre-network age there were some remarkable experiments in making more widely available the information stored in large national libraries, such as the book box scheme by which the National Library of Wales provided duplicate items to adult education classes in the 1930s (Baggs, 1997), these pale by comparison with the transformation wrought by the availability of digital and network technologies since 1980. Although librarians have always been fascinated by ways in which technology can reduce costs and streamline services, with Victorian librarians experimenting with early forms of duplication in cataloguing and the use of pneumatic tubes for transmission of book and other requests, the engagement of librarians and libraries with technology is nowadays more profound than ever.

‘The Battle of the Books’
The transformations are evident to anyone who haunts libraries, but, while the movement of the tectonic plates of scholarship is clearly evident, it is much more difficult to characterize and quantify the nature and impact of these changes. How can the impact of the digital transformations in the library be succinctly encapsulated? Do the new resources give the arts and humanities scholar unparalleled access to primary materials or does the deluge of data risk undermining the traditional relationship of arts and humanities research with books, archives and manuscripts? In 1993, at a time when many humanities scholars were still unfamiliar with the internet, Robin Alston gave a remarkable inaugural lecture as Director of the School of Library Archive and Information Studies at University College London entitled, after Swift’s satire on the 18th-century controversies about ancient and modern forms of learning, ‘The Battle of the Books’ (Alston, 1993). As Director of the Eighteenth-Century Short Title Catalogue project in the British Library from 1977, Alston, by supervising the creation of a machine-readable catalogue of printing in English-speaking countries to 1800, laid the foundations for
PART 1: DIGITAL TRANSFORMATIONS IN LIBRARIES, MUSEUMS & ARCHIVES

many of the online resources that are today most familiar to humanities scholars, such as Early English Books Online and Eighteenth-Century Collections Online. Nevertheless, Alston is far from an uncritical technophile, and in his inaugural lecture issued a warning against those who declared that:

the library has become less a nursery of knowledge than a play-room in which those with antiquarian obsessions indulge their fantasies and must be transformed into electronic warehouses invisibly connected to all other warehouses; that there can be no real progress until every scrap of paper has been digitized and rendered tractable to electronic manipulation and distribution via the telecommunication networks being built in space.  

(Alston, 1993)

Such an approach, Alston declared, ‘may be a fine prospect for faculties of engineering, and the makers of electronic devices on which the librarian is increasingly dependent, but it might have serious consequences for a balanced view of who we are, where we came from, and where we are going’. Alston stressed that ‘the reduction of analogue information to digital form, and the transformations which are then possible, is without doubt a process as fascinating as anything in the history of technology’, but felt that the view that libraries and archives simply comprise ‘so much information’, ignoring the complex bibliographical and physical character of the historical materials in research library collections, posed immense intellectual and cultural dangers. For Alston, these concerns related not simply to obvious logistical issues such as cost, preservation and copyright, but also to the way in which fundamental questions about the value of digitization had not been fully discussed:

Of course it is possible to digitize and index the contents of all the world’s important libraries and archives. In the 1930s Eugene Power demonstrated that it was possible to persuade libraries that what they needed was microfilm and created that juggernaut of unselective micropublishing University Microfilms. The question we must answer is, who could benefit from such a colossal enterprise? Commerce or knowledge? Are libraries in the control of visionaries or are they in the control of irresistible economic forces which we ignore at our peril?  

(Alston, 1993)

The real danger, as Alston saw it, was to the books themselves, and to the engagement with the physical culture of the book that is at the heart of much scholarly thinking in the humanities. To avoid these perils, Alton foresaw the need for:

enlightened and imaginative librarians able to develop within a hostile political environment a model which can adapt to the evolving needs of research in all disciplines. Those needs will, within a decade, include access to information in a wide variety of databases, electronic archives of images and sounds, as well as the cultural inheritance in print and manuscript, most of which will, I am certain, remain in its present form for the foreseeable future. One
consequence of this is the self-evident need for librarians in the future to develop both ancient and modern skills. The notion that knowledge can prosper by creating vast knowledge warehouses based on the hypermarket model – you can buy it if you can find it – is sheer fantasy as well as being intellectually suspect. (Alston, 1993)

Mass digitization initiatives
Alston’s 1993 warning against the hypermarket model seems particularly prescient today in the light of the subsequent development of Google Books and other cloud-based repositories such as the HathiTrust (Malpas, 2011). Alston’s lecture is an impassioned plea for a strategic approach to digitization, informed by awareness of the structure, history and physical characteristics of the collections being digitized. In his concern about the risks posed by commercial interests to the public character of the cultural heritage of libraries, books and archives, Alston was anticipating by many years the recent articles by Robert Darnton, the President of Harvard University Libraries, drawing attention to the dangers posed by the development of Google Books. Like Alston, Darnton’s concern is not simply that ‘Google will enjoy what can only be called a monopoly – a monopoly of a new kind, not of railroads or steel but of access to information’ (Darnton, 2009a) but also that the Google project is dominated by technical experts with little involvement of librarians or bibliographers, who would have avoided the ‘missing pages, botched images, faulty editions, omitted artwork, censoring, and misconceived cataloging that mar Google’s enterprise’ (Darnton, 2009b). The recent call by Darnton (2010a, 2010b) for the creation of a Digital Public Library of America looks very much like a belated attempt to shut a stable door behind a horse whose escape was noted 17 years ago by Robin Alston. It is striking how many of the anxieties described by Robin Alston in his inaugural lecture remain current. In particular, the question most forcefully posed by Alston, of the need to consider the purpose and value of digitization before embarking on large-scale investment and the creation of digital libraries, remains inescapable, and is still rarely addressed. Alston’s comment that ‘A close examination of the products currently available in electronic form might well suggest to an unprejudiced mind that we have a wonderful solution in search of problems’ (1993) sadly perhaps still holds true today. In 1993, Alston was unable to discern any immediately tangible benefits in digitization work in libraries.

Revisiting the ‘Battle of the Books’ after nearly 20 years, which have seen major funding programmes, the development of new technical facilities and the appearance of new services such as Google, it is surprising how far still lack tools or analytical approaches that can effectively measure or model the impact of digital resources on arts and humanities research. While statistics collected by libraries document major changes in provision since 1993, they give little sense of the extent to which these developments have reshaped or enhanced humanities scholarship or developed new audiences. The most striking shift in the nature of library provision since 1993 is that the bulk of the acquisition budgets of university libraries in the UK are today used for the purchase of access to electronic resources. In 2007–8, over 40% of expenditure on information
service provision in British universities was on electronic resources (if journals purchased in both print and electronic forms are included in the figure) (British Library, 2010, 9), and for the research-intensive Russell Group universities, the proportion was considerably higher. Notoriously, of course, the nature of the expensive and complex licences for online journal bundles means that commercial publishers have in effect hijacked large parts of university library budgets (so that recently the University of California threatened to cancel its subscription to *Nature* and to boycott peer reviewing activities for the journal) (Howard, 2011). Nevertheless, it is clear that there has been a major and radical shift in provision.

Given the nature of the large bundled journal subscriptions, it is difficult to disentangle the extent to which this pattern of expenditure reflects shifts in provision for the humanities, but changes in user behaviour are more easily documented. A survey in 2006 by Stephen Brown and Mark Greengrass (2010, 13) found that over 60% of a sample of arts and humanities researchers regarded digital resources as essential to their research, with 89% using the web on a daily basis. Focus groups conducted by Brown and Greengrass found that researchers considered that the greater speed with which material could be located and accessed had transformed the way in which they conducted research. Researchers felt that, as a result of these developments, new scholarship and ideas were circulating more quickly in the humanities, although they were vague as to the type of innovative thinking generated by these more active dialogues. Unfortunately, studies such as those by Brown and Greengrass or by Claire Warwick et al. (2006) are concerned with general patterns of humanities use of online resources, and do not tend to distinguish between specialist digital humanities resources and the more generic packages procured through libraries such as Early English Books Online, but there are indications in the study by Brown and Greengrass that these generic packages are the most important to library users. The British Library’s Online Public Access Catalogue was first made available online in 1997. By 1999, it was receiving more than five million hits a year. In 2009–10, the overall British Library website received more than 74 million hits. This represents a major widening of access to the library: in 2010 there were over 500,000 on-site visits to the British Library, a tiny amount compared with the huge figures for online access. In 1979, the British Library Lending Division received over two million requests a year, about 80% of which were satisfied, usually by manual means. In 2009–10, over ten million items were supplied to users by the British Library, the vast majority of these online (British Library, 1979–2010).

While figures like these illustrate the transformation that has occurred in library use and provision over a short time, they do little more than indicate that the infrastructure of humanities scholarship has changed. The usage statistics compiled by libraries tend to cover all disciplines, and it is difficult to disaggregate figures for humanities users from them. Moreover, the patterns of change are complex, and it is difficult to reach straightforward conclusions about the impact on humanities scholarship. For example, the digitized version of the *Codex Sinaiticus* was made available by the British Library in July 2009, and received over one and a half million hits during the following few months
(British Library, 1979–2010). Some of these hits were doubtless from biblical scholars but the vast majority were apparently from curious members of the public. While the digitization of the Codex Sinaiticus has demonstrably enhanced public engagement with the original manuscript, it is more difficult to quantify its impact on scholarship in the field. Moreover, simply to focus on statistics for the production of digital images as a measure of the impact of digitization is misleading, since the process of digitization covers a variety of activities within the library, and these different activities have had varied purposes, audiences and impacts. Digitization has not been restricted to rare books and manuscripts. JSTOR is the digitization of a large back catalogue of periodicals. Likewise, the conversion of printed and card catalogues to a machine readable form is also a form of digitization (indeed, in the British Library, the first attempt to convert the printed British Museum catalogue involved the early use of a form of optical character recognition (OCR)). It is striking that humanities scholars seem to take the availability of automated catalogues and searchable back runs of periodicals for granted, and it could be argued that to date these represent the major impact of digitization on the engagement of humanities scholars with libraries.

**The impact of digital transformations in the library**

In seeking to find a framework to analyse and explore the nature and character of these changes, it is perhaps worth returning to the idea at the heart of Manguel’s reflections in his *Library at Night*, namely the library as space and the way in which technology changes this space. The digitization of catalogues is a powerful reminder of how changes in libraries reflect changes in configuration in space. To digitize a catalogue and make it internationally available fundamentally alters the relationship of a library to the outside world. The catalogue, instead of being a map of the enclosed restricted space of the individual library, becomes instead a means by which this space is linked to the outside world, and, through such links to other libraries, the isolated library suddenly becomes part of a network. This important theme was explored by Lorcan Dempsey (1999) in a reflection on digital information spaces, which although it was published more than ten years ago is still worth re-reading. Dempsey emphasized how the traditional library was place-based, comprising the management of multiple individual physical repositories (the strict division of the historic British Library into departments based on particular media exemplifies this). By contrast, suggested Dempsey (1999, 54):

> In digital information spaces, there will continue to be repositories of information but the emphasis shifts to the flows between them and between them and their users. These repositories may contain metadata – catalogues and other data which assists in the discovery, use and exploitation of resources – or resources themselves. In this environment, the activities of discovery, locate, request, and deliver, currently carried out in multiple incompatible circuits need to be brought into a common framework of communicating applications.
PART 1: DIGITAL TRANSFORMATIONS IN LIBRARIES, MUSEUMS & ARCHIVES

The implications of this spatial readjustment of the library described by Dempsey are still barely worked through, particularly in research and university libraries. As well as different flows of information, the patterns in which information is held have also fundamentally shifted. The traditional library was primarily concerned with pulling in the widest range of information possible, and this is still an important feature in a digital environment. However, increasingly libraries are concerned with the curation of local information and pushing that information out to the network. This is a relatively new role, and one whose implications for scholarly work practices are still not yet wholly evident. One expression of this new concern is the institutional repository, whose origins lay in an attempt to undermine the monopolies of publishing houses on the dissemination of research produced in universities, but which is being increasingly used by universities as a means of managing and curating information about the research output of the university. For university libraries — and indeed universities — to take a direct role in managing and making available the research produced in the university has radical implications for scholarly practice that have only just begun to be explored. Two influential American scholars and commentators on libraries, Robert Darnton (2010b) and Jerome McGann (2010), have recently emphasized the profound impact on the academic career structures of the collapse in sales of academic monographs by American university presses, causing publishers to withdraw completely from certain subject areas. Possibly the increasing availability of doctoral theses through institutional repositories may make the production of such monographs redundant. Perhaps repositories may ultimately become the means by which niche academic writing is eventually disseminated (MacMahon, 2009). It is difficult to say at present, but certainly the humble institutional repository, by involving universities directly in the curation and dissemination of the research work of members of the university, has the potential completely to reconfigure the spaces of scholarly production and communication.

In concluding his 1999 article, Dempsey pointed out the paradox that, at the time the British Library was undertaking its first experimental programme of digital and network activities, it was also constructing at St Pancras ‘one of the most significant library “places” the world will have seen, a building in which the main objective is to create “an easy commerce between the lone scholar and the huge building mass required to house the collections, all the fellow (rival?) researchers and the general public”’ (Dempsey, 1999, 58). Dempsey noted the need for the new British Library building ‘to organise a social and learning space within the emerging global space of flows’. The success of the British Library building says much about the continuing demand for a physical space for work, reflection and investigation within the global space of flows, and the way in which the restaurants and cafés of the British Library have accommodated what St John Wilson called ‘the wandering scholars of the laptop’ suggest that the British Library building, designed long before the internet, is impressively adaptable to these new flows.

**The transformation of library space**

The growth of network services has led to a reconfiguration of physical space in the library,
not its disappearance. While online services have grown in recent years, library building has also boomed. Partly this reflects the fact that the production of printed books has continued unabated: Robert Darnton (2010b) calculates that there will soon be a million new titles published worldwide each year, adding that ‘I have been invited to so many conferences on the “Death of the Book” that I suspect it is very much alive’. Likewise, far from becoming more marginal places, university libraries have been under pressure to increase their opening hours, with many British university libraries now staying open into the small hours or even, in a few cases, 24 hours. Although principally a library designed mainly to serve scientists who make increasingly less use of printed resources, nevertheless demand by Imperial’s students for a flexible study space led to an 11 million pound refurbishment of Imperial College’s Central Library, which opened in 2008, and has proved immensely popular with students. Imperial College is not alone in its investment in new library buildings. The past few years has seen the opening of new university library buildings in Leicester, Sheffield, Coventry, Glasgow Caledonian University and elsewhere, while major new developments are currently under way in, for example, Aberdeen and Manchester (Designing Libraries, 2011). However, many of these library buildings are radically different from those that were constructed during the period of university expansion in the 1960s. Features of these developments include the integration of many different types of information, with online provision being given particular prominence. These libraries are just as much social hubs, for meeting, talking, sharing information, as they are places of study. The coffee bars are as important as the quiet study rooms – the buildings are highly zoned. Correspondingly, many of these new forms of library also provide access to other university services, particularly student services. Above all, librarians are less evident as gatekeepers. This partly reflects the use of self-service technologies, but also the reconfiguration of support services, which have been placed in help desks or reconfigured as roving assistants.

Some of these developments, such as the new emphasis on group work, are driven more by pedagogical than technical considerations. However, it is clear that a major force in shaping these new libraries and learning spaces was the demand for online access in a variety of forms alongside more conventional library resources. The corporate university information services (whether converged or not) play as important part as the library in designing these new facilities. As Diana Oblinger has observed in a recent publication by Educause on new learning spaces:

Information technology has changed what we do and how we do it. It would be hard to identify a discipline in which IT is not a necessity. Collecting, analyzing, displaying, and disseminating knowledge typically involves IT. Retrieving information has become an IT function; students consider the Internet, not the library, their information universe. And, rather than trying to know everything, students and faculty rely on networks of peers and databases of information. What impact, if any, should this have on learning space design? Technology has also brought unique capabilities to learning. Whether by stimulating more interaction through the use of personal response systems or by videoconferencing with international experts, IT has altered learning spaces. (2006, 1.2)
PART 1: DIGITAL TRANSFORMATIONS IN LIBRARIES, MUSEUMS & ARCHIVES

It might be asked how far this new breed of library building is an attempt to save the concept of a library fatally wounded by the delivery of information services to student homes and staff desks, but the demand for designated learning spaces of this sort is strong and evident: in the case of Sheffield, Martin Lewis (2010, 176) has noted that ‘The biggest challenge faced by the Information Commons is that of its own success… Maximum occupancy is regularly achieved and the most common complaint from students is inability to find a seat, despite the delivery of a substantial net gain in overall study space capacity on the campus’.

Such new learning spaces have a varying relationship to the library. In some cases, such as the Learning Grid at Warwick University, the new learning space supplements the central library; in others, most notably the Saltire Centre at Glasgow University, the library has been transformed into a social hub for the entire campus. In many other cases, existing buildings have been remodelled and reshaped. A good example of the refurbishment of older library buildings to meet new needs is the Library and Information Centre at the University of Swansea. The main reading areas of an older building have been refurbished to provide access to banks of computers adjacent to a coffee shop and with various student service points close by. The immediate impression on entering the Swansea Library and Information Centre is of noise, interaction and social exchange. Research libraries have begun cautiously to import some of this thinking: the recent refurbishment of the North Reading Room of the National Library of Wales has introduced into a formal reference reading room from the 1930s group study areas, informal seating, extended provision of personal computers (PCs) and wireless access. These new library spaces are the most striking expression of the impact of digital technologies in the library. However, by focusing on these spaces, there is a danger of falling prey to older perspectives of the library, and to think of individual libraries as autonomous and disconnected spaces. It is in the flows of information between libraries and the way in which new connections are developed between them that the transformations in libraries are more evident. This requires a study of libraries based on national and international links. For larger countries such as France or the USA, this would necessitate a major study, but the transformational effects of digital technologies can be more readily discerned in small countries. The way in which, for example, the national libraries of Finland and Sweden enthusiastically pioneered the use of digital technologies in the 1980s and 1990s shows how digital and networking technologies can provide business, educational and cultural benefits in countries where a small population is dispersed and the physical geography makes face-to-face communication difficult. In Great Britain, this ‘small country’ model can be seen at work in the development of library provision in Scotland and Wales.

Case study: Library initiatives in Wales

Introduction and background

In order to get a sharper idea of the extent to which changes in libraries are reshaping research in the arts and humanities, Wales provides a very interesting case study.
Formidable barriers of physical geography and the resulting difficulties in communication have been major issues in the development of Wales since the middle ages, and the persistence of the Welsh language despite the intense competition from English is partly a reflection of the physical isolation of parts of Wales. Even today, travel between north and south Wales is very difficult, and improving links between the different parts of Wales in order to create a greater sense of shared nationhood has been a priority of the Welsh Government since its establishment (as the Welsh Assembly Government) in 1999. The communication difficulties in Wales have encouraged the development of decentralized and federal institutions. The University of Wales is perhaps the best known example of a federal national institution in Wales, while the Welsh Government has been at pains since its creation to ensure the establishment of an extended network of regional offices. While Cardiff became the home to such Welsh national institutions as the National Museum of Wales and the Registry of the University of Wales during the period before the First World War, it was not designated as the capital of Wales until 1955, and even then other towns such as Caernarfon vied for the honour. The pattern of cultural and administrative provision in Wales remains very dispersed. The National Library of Wales for example was established in the small western coastal town of Aberystwyth at the insistence of its major benefactor, Sir John Williams, who was anxious to see the library placed in a Welsh-speaking area. The National Library is thus located in an area that, as the present National Librarian of Wales Andrew Green (2007a) has commented, is ‘well populated with animals but less well endowed with humans. There are about 4.5m sheep in our hinterland, compared with only 180,000 people, and to make matters worse not all of the sheep are yet fully literate. We’re seventy miles from any major town or city.’

**Collaborative frameworks in Wales**

The need to transcend these physical barriers meant that the National Library and other Welsh institutions took an early interest in digitization. While this is readily understandable, what is striking about the Welsh experience is the extent to which it was underpinned by a strong collaborative framework. Much of this has been directly driven by the Welsh Government, which has since its creation consistently emphasized the need for reconfiguration and collaboration in Welsh public institutions. This led, for example, to the creation in 2004 of CyMAL as a policy division of the Welsh Government to foster collaboration between libraries, museums and archives (Atkinson and Kenser, 2004, 51–2). Among the development activities taken forward by CyMAL has been the creation of an integrated portal for Welsh libraries (http://library.wales.org) featuring CatCymru, which allows the cross-searching of all publicly accessible library catalogues in Wales. CyMAL has also facilitated the provision to all library members of online resources such as Newsbank. Another major focus of cooperation has been the Welsh Higher Education Libraries Forum (WHELF), comprising the librarians of higher education institutions in Wales together with the National Library (Atkinson and Kenser, 2004; www.whelf.ac.uk). These collaborative activities have been actively facilitated and promoted by the National Library, and the active role of the National Library perhaps helps explain why patterns of
library provision in Wales are less fragmented than in some other areas of the UK. The extent and effectiveness of collaboration among the Welsh libraries is very striking and suggests that strategically directed partnerships can be very important in developing successful digitization projects that achieve major impact.

One focus of Welsh collaborative activity has been joint procurement of commercially produced digital resources. The most striking digital divide at present is between members of universities, who have access to commercial journal bundles and such packages as Early English Books Online or British Library Newspapers, and the general public for whom the cost of access to such resources is prohibitively expensive. For most university-based researchers in the UK, the effectiveness of the agreements negotiated by the JISC for discounted access to commercially produced journal bundles and databases means that there is little recognition of the high cost of these resources. In the USA, there is no such national procurement, so that many smaller colleges find the cost of online journal packages prohibitive, leading to a digital divide between small colleges where the provision of electronic resources is patchy and larger universities, which have invested heavily. Although UK university researchers benefit from the national agreements brokered by JISC, the general public are still frequently excluded from access to the main online resources on which researchers rely. Like the National Library of Scotland, the National Library of Wales makes available to all its registered readers who have a Welsh postcode a portfolio of electronic resources including many of those most familiar to humanities researchers in universities. Although any member of the Welsh public can become a registered reader, Andrew Green (2007b) has expressed his ambition to ‘increase the size of this digital collection, and to extend its reach far beyond registered readers’. For Green, the establishment of such a service ‘will mean a radical rethink of what it means to be a Library member ... the result being a truly national online knowledge service available freely to every Welsh citizen’. Green points out the parallels between this vision and the model adopted by the National and University Library in Iceland, which provides free access to over 8000 full-text journals and databases for all Icelandic citizens. The creation of these ‘small country’ online national libraries may also be seen as prefiguring the recent discussion initiated by Robert Darnton (2010b) for the creation of a Digital Public Library of America. The recent Research Information Network (2011) report on e-journals has documented the various ways in which these have transformed scholarly practice, yet this revolution remains in many ways effectively confined to university researchers, and this knowledge is perhaps less widely accessible to other parts of society than it was in the print era. The initiatives adopted in Wales and other small countries to open up access to these online resources are a vital part of the way in which libraries are renegotiating their role as knowledge brokers in the online world.

These activities of the National Library of Wales have been complemented and extended in the higher education sphere by various initiatives of WHELF. Particularly notable was the creation of a shared Welsh university collection of e-books, which was launched in the autumn of 2007 with over 500 e-books on OCLC’s NetLibrary (Atkinson and Riley, 2009). The experience gathered through this shared e-book project will be
important in negotiating future access to new forms of e-books as these become more widely available. The value of a collaborative approach in ensuring a strategic approach to new services is also apparent in WHELF’s Welsh Repository Network project, which created an integrated network of repositories in all Welsh universities, enabling many advantages such as the development of a cross-searching facility and the development of a service for storing and harvesting electronic theses produced in Wales, which it is hoped will increase the visibility and usage of Welsh theses and give Wales a research advantage (Knowles, 2010). A more recent WHELF initiative has emulated the pioneering work of its Scottish opposite number, SCURL (Scottish Confederation of University and Research Libraries), in securing joint licences to e-journals to create a Scottish Higher Education Digital Library.

The digitization of Welsh cultural heritage
While on the one hand both WHELF and the National Library have collaborated in realigning audiences for digital resources through innovative procurement activities, perhaps more pertinent for the themes of this volume is the role of both WHELF and the National Library in digitizing historical collections in Welsh libraries. While the physical location of the National Library was certainly one incentive for its early interest in digitization, another was the very wide range of media held by the National Library. As well as books, manuscripts and maps, the National Library also houses the de facto national archive of Wales, its national photographic collection, its screen and sound archive, and its second largest art collection. Indeed, it was the presence of the screen and sound archive that provided the initial impetus for the National Library of Wales to experiment with digitization, with a project to create stills from films for the Welsh language television channel S4C (Green, 2007a). The Library took an early decision to use mainstream National Assembly funding to support its digitization programmes; Andrew Green points out that this contrasts with the British Library, which has never used core funding for digitization but relies on external funding and outsources major digitization operations. Green (2007a) adds, ‘We wanted to develop a body of internal practice in the art and science of digitisation: not just data capture, but the entire workflow, from selection to presentation. We built up small teams of scanners and metadata experts, and added others to contribute in other areas, such as preservation and online presentation.’ Because of this stress on integrated management of digitization process, including metadata, the Library pioneered in the UK the Metadata Encoding and Transmission Standard (METS) schema developed by the Library of Congress.

Much of the National Library’s early work in this sphere comprised a series of projects designed to promote awareness of the Welsh cultural heritage, which were administered through Culturenet Cymru, a limited company owned by the Library and dedicated to promoting online culture in Wales. Among the largest projects undertaken by Culturenet Cymru was Casglu’r Tlysau/Gathering the Jewels (www.gtj.org.uk), a People’s Network project containing over 30,000 images from over 140 libraries, museums and archives around Wales. The involvement of the three domains of libraries, museums and archives
mean that the subjective experience of using Casglu'r Tysau is immediately very different from that of a conventional library. The materials made available via Casglu'r Tysau range from a Welsh woman's hat of the 19th century, stuffed animals from West Wales and aerial photographs of hill forts to the first known recording of the Welsh national anthem and manuscripts by Iolo Morganwg that recorded his invented Druidic histories. Such a description might make Casglu'r Tysau sound like a digital equivalent of a rambling antiquarian museum, but the effect is far more integrated and powerful and it is not an exaggeration to describe it as a reinvention of Welsh national culture on the web. While there can be no question about the popular impact of Casglu'r Tysau, there is far less evidence of its impact on scholarly and research activities. Casglu'r Tysau makes available full facsimiles of some manuscripts but, like its Scottish counterpart, SCRAM (Scottish Cultural Resources Access Network), its focus is on supporting learning and in increasing general awareness of the resources of Welsh museums, libraries and archives.

This community emphasis of the National Libraries digitization programme was taken a stage further by another Culturenet Cymru project, Your Past Their Future (www.pyf-wales.com). This project was linked to a touring exhibition commemorating memories of the Second World War, and recorded the recollections of Welsh people about the war. Items from their personal collections, such as photographs, ration books and letters from the front, were digitized and added to the website to form a virtual archive of the Welsh experience of the war. Again, like Casglu'r Tysau, the main function of this activity is in the area of community engagement, and its direct value in arts and humanities research appears comparatively limited. Projects such as Your Past Their Future and Casglu'r Tysau also have a major function, in a devolved administration, of assisting in nation building. This aspect is even more strongly apparent in Casglad y Werin Cymru/People's Collection Wales, which was launched in August 2010 as a result of a specific commitment in ‘One Wales’, the agreement governing the coalition of the Labour Party and Plaid Cymru in the Welsh Assembly Government at the time. Casglad y Werin Cymru offers seamless access to digital images from the National Library of Wales, the National Museum of Wales, the Royal Commission on the Ancient and Historical Monuments of Wales as well as to Casglu'r Tysau. Cross-searching of this type between museum, library and other collections can potentially be very beneficial for research purposes, but the structure of Casglad y Werin Cymru is much more geared to popular access and engagement. Members of the public can add their own images, create exhibitions that they can share online, create their own family stories, and map out walks and tours. Notwithstanding the strongly popular emphasis of the site, usage appears to be comparatively limited: a visit to the site in March 2011, 18 months after the launch of the site, showed that it had attracted only 1249 contributors.

A holistic approach: management of the digital life cycle
While these figures may seem to suggest that the impact of Casglad y Werin Cymru has so far been comparatively limited, it is important to note that it forms part of an integrated approach to digitization by the National Library of Wales. Such holistic management of digitization and the storage of digital data is important in achieving the
maximum use and impact from digital resources. The wide range of media within the National Library posed pressing issues with the management of digital data from an early stage in the library’s digitization programme. The National Library had to address not only the storage of images of historic books and manuscripts but also the management of large quantities of digital recordings of Welsh-language television programmes and increasingly the deposit of born-digital information. Such born-digital acquisitions can be very large; on the closure of a Cardiff media company in 2006, the Library acquired a shared drive containing the records of the company, which contained 26.2 gigabytes (Gb) of data (Robson, 2010). It is also usually necessary to transfer such born-digital data from physical carriers that may already be faulty by the time they reach the Library; in a deposit of compact discs (CDs) by Cliff McLucas and Brith Gof Theatre Company, 20% of the CDs from McLucas and 60% of the CDs from Brith Gof could not be read on arrival in the Library (Robson, 2010). In order to manage the ingest, metadata, storage and access of this wide range of materials, the National Library implemented a digital asset management system (DAMS) using the Vital institutional repository system (www.vitals.com/products/vital), which is based on the open source FEDORA (Flexible Extensible Digital Object Repository Architecture) software (Bevan and Robson, 2008).

The quantity and variety of material already handled by the National Library’s DAMS is extremely impressive. By February 2009, the DAMS was holding for example over 800,000 images of 186,000 wills dating from between 1543 and 1858, amounting to some 2.5 terabytes (Tb) (Robson, 2009). At the same time, over 1000 television programmes were being automatically recorded and ingested into the system. From the point of view of the researcher, the value of the DAMS lies in the way in which it integrates digital objects with the library’s conventional collections. Geoff Charles (1909–2002) was a newspaper photographer who over 50 years produced a vivid and distinctive portrait of Welsh life, and his archive of 120,000 photographs is now one of the National Library’s treasures. In the late 1990s, it was noticed that some of the cellulose acetate negatives in Charles’ archive were suffering from ‘vinegar syndrome’, and a programme to digitize the affected negatives was established so that the original negatives could be frozen to stabilize their condition. Over 14,000 images from the Geoff Charles archive were digitized, producing 2294 images and over 16,000 FEDORA objects (Bevan and Robson, 2008). The fact that these objects are held within the DAMS means that they can be made available in a variety of ways: the Charles archive is one of the prominent components of Casgliad y Werin Cymru; it features in the ‘digital mirror’ section of the Library’s website; it supports the National Library’s contribution to the Flickr Commons site, which allows users to comment on and tag photographs from major museums, libraries and archives (www.flickr.com/photos/llgc); and a search on the Library’s catalogue of terms such as Wrexham will also produce images by Charles, integrated into the catalogue record. The use of a DAMS to manage the Geoff Charles images is a small but important example of how the impact of digitization activities can be maximized. By holding the images in a digital repository, they can be made available in a variety of contexts, meaning that they can be used to address the needs of different audiences.
PART 1: DIGITAL TRANSFORMATIONS IN LIBRARIES, MUSEUMS & ARCHIVES

Public engagement and access with the Geoff Charles archive can be promoted through sites such as Cysgod y Werin Cymru, while at the same time researchers may use the archive in a more structured fashion through the Library’s catalogue.

Among the key lessons that Glen Robson, the manager of the DAMS at the National Library, has drawn from experience thus far is that digitization should be organized around ‘workflows not projects’ (Robson, 2010). The benefits of integrating the digitization of a particular archive into a wider workflow are already apparent in the case of the Geoff Charles archive. Treating digitization activities as a series of discrete projects requires customization in respect of such elements as metadata or web presentation for each project. A focus on workflows means that much the same approach can be used for handling the digitization of wills as was used for the Geoff Charles archive, potentially reducing costs and increasing impact. However, researchers in some disciplines may feel uneasy about the stress on generic workflows. While librarians tend to be doubtful about the value of ‘boutique digitization’, arts and humanities researchers are often much more interested in comprehensive and scholarly editions of particular manuscripts and books.

Another of the early fruits of the National Library’s digital work was the creation of a ‘digital mirror’ comprising high quality digital facsimiles of some of the treasures of the library, such as an illustrated 13th-century copy of Laws of Hywel Dda, the first book printed in Welsh from 1546, and Dylan Thomas’s map of Llareggub. Unlike the British Library, the National Library has consistently made complete facsimiles of the treasures available, so that the digital mirror in effect forms a small online manuscript library. Hitherto, the ‘digital mirror’ has been presented as a separate section of the National Library’s website. An advantage of the DAMS is that the material in the ‘digital mirror’ will be presented in a different format in future and more closely integrated with other digital collections.

Manuscript digitization

Among the material presented on the ‘digital mirror’ are images of manuscripts of poems attributed to the celebrated 14th-century Welsh poet Dafydd ap Gwilym. These images were supplied for an interdisciplinary project at Swansea University directed by Professor Dafydd Johnston, which produced a comprehensive online edition of the poems of Dafydd ap Gwilym and was launched in 2007 (www.dafyddapgwilym.net). By presenting the manuscript evidence for these celebrated Welsh poems online, Dafydd Johnston and his team were able to show how readings for many well known lines were questionable. The online edition of Dafydd ap Gwilym is, from a scholarly point of view, an exemplar of how an online edition can deliver concrete research results. However, there can be no doubt that it falls into the category of ‘boutique digitization’. Another of the National Library’s great treasures, the early manuscript of Chaucer’s Canterbury Tales, Peniarth MS 392D, was also the subject of a similar digital edition by Estelle Stubbs (www.sd-editions.com/hengwrt/index.html). Stubbs’ work helped to contribute to the proposed identification of the London scrivener Adam Pynkhurst as the scribe of the manuscript, and has proved very influential on scholarship in the field. However, the CD edited by
Stubbs is another example of ‘boutique digitization’. While from the point of view of the creation of a digital library, the advice offered by Glen Robson to focus on workflows not projects has a great deal to commend it; the stress in much humanities scholarship on the importance of the edition means that much digitization work in the arts and humanities will inherently need to be project-based if it is to achieve scholarly impact.

**Mass digitization in context**

There is potentially a tension between the requirement of scholars for project-based work and the need for libraries to emphasize workflows. The importance for librarians of workflows and a strategic approach to digitization has becoming more pressing in recent years. This is partly due to the arrival of the Google Books programme at the end of 2004. If projects like Dafydd ap Gwilym and the Hengwrt Chaucer represented ‘boutique digitization’, then the arrival of Google Books represented what Andrew Green (2007a) (in an echo of Robin Alston’s comment about hypermarkets of knowledge) has called the IKEA era of digitization. According to Green, since the advent of Google Books ‘most libraries have aspired to – if they haven’t always found the means to realise – much larger-scale, mass production projects . . . Selection is simplified: whole collections can be treated, without worrying about making choices. Volume sometimes wins over quality, speed over care; but the end result can have more utility’ (Green, 2007a). A more particularly pressing problem for national libraries in countries such as Wales is the way in which Google Books effectively excludes the literature of small countries. The library partners in Google Books are dominated by large American university libraries. The only British partner is the University of Oxford. While there are libraries from other major European language areas such as Germany and France (and one university from Japan), there are no representatives of countries such as Denmark, The Netherlands, Finland or any Balkan countries. This means that Google Books gives an inherently distorted view of the printed inheritance, so that many early Welsh books (including apparently the first Welsh Bible of 1588 that was critical in ensuring the survival of the Welsh language) are either completely invisible in Google Books or only present through later reprints and facsimiles.

Of course, while Google Books has increased awareness of ‘big digitization’, it is by no means the only approach to large-scale digitization of libraries, as Andrew Green pointed out in an important lecture to the Digital Resources in the Humanities and Arts Conference (2009). The Open Content Alliance was established in 2004 in response to Google Books, but has not been active in scanning of material since the withdrawal of Microsoft from the alliance in 2008. The Internet Archive has been active as a repository of open access material and the Open Library in many respects provides more extensive coverage than Google Books. Other libraries and archives have preferred to work with commercial partners in developing large-scale digital projects. The National Archives has used commercial partners to digitize British census returns, recouping costs by charging users. Likewise, the British Library is using a commercial partner in the digitization of out of copyright newspapers.

For Andrew Green, such a commercial approach presents great problems:
Those who advocate this approach point to its potential to make rapid progress, and to provide sustainability to a continuing digitisation programme – assuming, of course, that the commercial model itself holds water. The major disadvantage is the loss, at least for a considerable period, of either the wide public access that lies at the heart of the mission and indeed justification of national public institutions, or the free use of information that is also one of their central characteristics – or both. Libraries then become indistinguishable in their functions from publishers. An alternative is to view both the creation and provision of digitised knowledge as a public good. Public libraries were established and still flourish as a means of ensuring that all members of society, irrespective of their circumstances, could have access to all published knowledge. Traditionally their method of achieving this goal was to offer a place where any citizen could read and usually borrow any printed publication. The 21st century equivalent is surely to offer citizens free access, where possible online, to publications in digital form. It follows that, in the case of those publications already in print, libraries – in particular, I would suggest, national libraries, that have a duty to preserve and give access to their countries’ published output – should do their best to arrange for their online public accessibility.

(Green, 2009)

Green pointed to the ambition of the national libraries of countries such as New Zealand, Slovakia, Norway, Finland and The Netherlands to digitize as much of their collections as possible, and pointed to this approach as a means of achieving a network level of impact for the libraries and literature of those small countries ‘of little interest to the global giants of Google’. The National Library of Wales has committed itself to a similar vision for Wales. The achievement of this vision will inevitably be a slow process, but a major first step is Cylchgronau Cymru Ar-kin/Welsh Journals Online, which was launched in February 2009. Cylchgronau Cymru makes available online 52 periodicals from Wales, including titles in both Welsh and English. It contains 400,000 pages altogether, with 180,000 pages in Welsh, representing the largest single corpus in the language available on the web (Locock, 2009, 5). While the material in Cylchgronau Cymru is exposed to Google, fully structured metadata in a MArchine-Readable Cataloguing (MARC) 21 format was also prepared for each article in the journals in the project, producing a very rich bibliographical record, and the provision of this metadata proved to be one of the most demanding aspects of the project.

Two other features of the project are of particular note. The first is that most of the journals date from the 20th century and consequently most of the articles in them were in copyright. Copyright agreement was obtained on a case-by-case basis. This occasionally proved controversial (Meredith, 2008) but agreement was obtained for the overwhelming majority of articles, although, since no payment was offered, works by Dylan Thomas, Robert Graves and R. S. Thomas are not included in the online resource. (The project manager, Martin Locock (2009, 5), notes that given that the cost of web publication was about £2 per page, the payment of even minimal copyright fees would transform the economics of mass digitization.) The result of this work in obtaining clearance for the reproduction of work in copyright is that Cylchgronau Cymru is an addition, rather than a
rival to, Google Books, which does not fully reproduce material in copyright. The second notable feature of the project is again its collaborative character. The project was not simply a National Library project, but was undertaken under the aegis of WHELF, and indeed the active support and involvement of the WHELF libraries was essential not only in securing the necessary access and agreements for the digitization but also in ensuring funding under the JISC capital programme. Some funding for Cylchgronau Cymru was also provided by the Welsh Government, and this paved the way for additional funding through its Strategic Capital Investment. Found the lion’s share of the funding for a second larger project, now under way, to digitize newspapers and periodicals from the 19th century (Green, 2009).

Projects such as Cylchgronau Cymru have been carefully designed in order to make a strategic impact that both ameliorates the cultural issues raised by large resources such as Google Books and provides structured access to core elements of the Welsh printed heritage. This strategic approach is pursued collaboratively through bodies such as WHELF with the aim of creating an integrated and comprehensive resource that Andrew Green (2007a) has designated ‘The Theatre of Memory’. More importantly, such initiatives will themselves ultimately feed through to larger resources. The National Library is among the institutions contributing to the Europeana programme, a web portal that will bring together digital collections from a variety of institutions across Europe (www.europeana.eu/portal/index.html). Currently, the National Library of Wales is contributing content through the eContentPlus Programme under the theme of ‘travel’, and is digitizing 574 drawing volumes, containing over 35,000 drawings and watercolours, from the amateur sketches of wealthy travellers to the field sketchbooks of professional artists (http://nlwales.blogspot.com/2010/10/contributing-to-europeana-web-portal.html).

Recent developments in digitization in Wales provide some strong pointers as to how the impact of digital resources can be maximized. One key lesson appears to be the importance of what Andrew Green called a ‘whole process approach’, in which digitization is seen as ‘an integrated series of operations, starting with selection and preservation, moving through image and data capture and metadata addition, and ending with presentation’ (Green, 2007a). In this context, impact is enhanced if projects are designed to make a strategic impact in terms of the current and developing landscape of digital provision. In successfully achieving this, a strong collaborative and cooperative framework is desirable. Close alignment with the funding priorities of government and other bodies is also essential in securing financial resources for these programmes. The Welsh record is particularly impressive in terms of its success in securing financial support for digital programmes. While questions of cost occasionally loom, as with copyright payments for authors like Dylan Thomas or R. S. Thomas in Cylchgronau Cymru, they are not particularly prominent. This is partly because the integration of digital processes into overall workflows in the library has proved extremely cost-effective (as is evident from Lcock’s calculation (2009, 5) that the web publication cost of Cylchgronau Cymru is about £2 per page). While Lcock notes that the creation of metadata can be very expensive, it does not appear that the cost is much greater than
PART 1: DIGITAL TRANSFORMATIONS IN LIBRARIES, MUSEUMS & ARCHIVES

with conventional library materials. The success of the Welsh programme in securing financial resources to date also reflects to a large degree the strong strategic alignment of the programme, and particularly the way it links to Welsh Government priorities. The extent to which the current crisis of funding in the public sector will affect the programme in the future remains to be seen.

Supporting digital scholarship in Wales

However, the stress of the National Library and WH ELF on ‘big digitization’ raises questions as to the continuing role of scholarly projects such as the Dafydd ap Gwilym and Hengwrt Chaucer editions. Where do such scholarly resources fit into this picture? Has the National Library succumbed to the dangers outlined by Robin Alston in 1993 and decided to focus on the creation of a Welsh bibliographical hypermarket to the exclusion of the smaller scale scholarly investigation of the bibliographical and textual characteristics of manuscripts and rare books? Andrew Green has stated that he believes that small-scale digitization has an assured future – possibly not as end in itself but simply as a natural by-product of another process, such as an academic research project (Green, 2009). However, it is clear that he sees such activities as remaining on a cottage scale, and therefore of limited impact. Indeed, it seems that he sees the future of cyberscholarship as lying more in the exploration of the larger data sets produced by bodies such as the National Library:

In future, as techniques for searching, text mining, document recognition and automated translation become more sophisticated, as the semantic web develops and as cyberscholarship becomes more common, these great reservoirs of text will yield new knowledge, and perhaps even generate new research fields. As interest in this kind of specialist analysis grows it will ask complex questions about the digitised material itself: how, for example, might it be possible to extract consistent and structured information about places, people or dates out of a large mass of unstructured and heterogeneous data; how, in the absence of detailed structured metadata, might relevant items be recalled or grouped successfully; how OCR techniques might be improved or supplemented to achieve greater accuracy.

(Green, 2009)

While Green paints a picture of many enticing research vistas, there are many other areas of humanities research that might require different digital techniques. Paleographers, for example, might be more concerned with the way in which different scribes form their letters; art historians might wish to compare brush techniques in paintings; landscape historians might wish to compare and overlay details in different maps. These all require specialist studies that inherently must frequently comprise projects rather than workflows, and which sometimes it can be very difficult to integrate into workflows.

It is here perhaps we sharply encounter one of the main issues identified by Robin Alston in ‘The Battle of the Books’. ‘Big digitization’ projects treat the contents of books as so much information; the concern of humanities scholars is frequently with the
historical context of that information. For a palaeographer, an art historian, a textual editor or a historian the information cannot be separated from the medium in which it occurs. It is fundamentally affected by the physical character of the manuscript, book, painting, print or recording in which the information occurs, and for many humanities scholars it is the study of the physical characteristics of that medium that are the focus of their research. As projects like the Dafydd ap Gwilym or Hengwrt Chaucer editions illustrate, part of the attraction of digital technologies is the ability minutely to explore in new ways the characteristics of manuscripts, printed books and other historical artefacts. It was for this reason that Alston so forcefully reminded his audience that traditional bibliographical skills would continue to be important in the new technological era. Yet it is difficult to see how the approach of the digital humanities, which emphasizes this close examination of texts, manuscripts and books, can be scaled up to produce the impact achieved by ‘big digitization’.

Robin Alston felt that this reflected the inherent difficulty of measuring the impact of humanities research, asking:

What useful purpose is served by disproving the existence of Phalaris? Or that Athens could not have been supplied with grain from the Black Sea because the Bosphorus flows at an average of 5 knots and their best sailing ships could only just manage that speed, so that while they could come down they had no way of going up? (Alston, 1993)

The success of the National Library of Wales and its partners in convincing the Welsh Government of the value of spending three million pounds on digitizing 19th-century Welsh periodicals and newspapers suggests that it is possible to convince government and funding bodies of the importance of such research, and that the secret of success in achieving this lies partly in a strategic alignment with government priorities, an approach which may come more naturally to librarians than academic researchers. Alston imagined that, in the new era of libraries, the need for librarians who could mediate between the new technologies and the older forms of bibliographic understanding would be greater than ever. Certainly it does seem that, in order to achieve greater impact in the digital sphere, it is essential for academic researchers to forge new, closer methods of working with librarians and curators. This reflects the view expressed by Andrew Green on the future role of librarians who he feels will hold a unique set of functions in the new digital economy of knowledge, in particular by acting as a mediator between existing ‘stored culture’ and the re-creation and production of new living culture and knowledge (Green, 2007b).

Conclusion
Reading in his library late at night, bathed in pools of light and comforted by the familiar smells of his books, Alberto Manguel is seized by an elegiac mood. The libraries he loves seem to him doomed, and his book is to a large extent a lament for a remarkable human
PART 1: DIGITAL TRANSFORMATIONS IN LIBRARIES, MUSEUMS & ARCHIVES

institution that he feels will soon disappear. The danger signs are certainly there, and were there when Robin Alston described the new Battle of the Books in 1993. It is a common fear, described again by Andrew Green (2007b):

An apocalyptic forecast might run like this. Libraries and archives have no future except as museums of the written word. Future generations will get their knowledge online and direct, either free from its originators or on payment from giant aggregators like the fictional ‘Googlezon’. Even what libraries used to call ‘special collections’, historic or rare publications and unique archives, will have been electronically cloned and networked long ago by commercial organisations.

Our reading rooms will empty. Even when it is libraries that provide access to online information, its users will be unaware of the fact. Skills for which librarians and archivists used to be valued, such as metadata creation, will have been undermined by further improvements in search engine technology. Long before the bicentenary of the National Library of Wales in 2107 the last reader will have been escorted out, and the building sealed, preserved like the statue of Ozymandias in the sands to remind travellers of the absurdity of monumental ambition.

Just as Alston argued that librarians have a special role to play in helping to ensure that the Battle of the Books does not lead to their destruction in favour of cloud repositories, so Green suggests that libraries can lead a process of reinvention and re-creation that reinterprets such traditional forms of knowledge as bibliography or codicology in a digital world, and can help usher in new forms of public culture. The many recent developments in libraries, from the creation of new forms of library space to the advent of major digital projects such as Cylchgronau Cymru and Europeana, shows that this process of reinvention is proceeding apace, and that, despite initial appearances in some places, such as Wales, the Battle of the Books is being won. In order to achieve similar impact and to play their part in the Battle of the Books, humanities academics need to forge closer alliances with librarians and to show an equal willingness to develop new working relationships and public roles.
The digital museum

Claire Hudson

Introduction
UK museums have experienced an interesting and sometimes frustrating journey since the arrival of affordable digitization technology. Until very recently, the impact and value of museum collections, especially for the research community, has not been fully recognized, and this has tended to hamper the speed and extent of museum digitization projects. Gradually, museums have not only generated a significant body of digital content themselves but have also become partners in higher-education-funded research projects with digital outputs, encouraging wider awareness not only of the value of their collections but also of the expertise they can contribute to the creation of research resources.

Despite this, there is still a great deal of unexploited potential. A recent study by the Comité des Sages (2001) on the online accessibility of Europe’s digital cultural heritage identifies the museum sector as an area requiring greater investment. The physical characteristics of many museum artefacts have also tended to inhibit digitization. Frail materials, large objects, collections housed in off-site stores, or items of costume that need three-dimensional mounting to convey a true picture of their purpose and design all present time-consuming challenges for digitization that add to project costs.

This chapter will present a case study based on the experiences of digitization at one specific organization, my own, followed by ‘the bigger picture’, some more generalized perspectives on the implementation of digitization within UK museums as a whole.

Case study: A learning journey
New Opportunities Fund – the NOF-Digitise programme

Introduction
The V&A Museum’s Department of Theatre and Performance is by no means a leader in the field of digitization but has been a very active player, with a track record stretching over ten years – almost back to prehistory in digitization terms.

In common with many UK heritage organizations, the journey started in the late 1990s when the V&A Department of Theatre and Performance was still known as the Theatre Museum, a branch of the V&A, located in its own premises in London’s Covent Garden. Although dependent on the parent organization for many of its support services (IT, human resources (HR), conservation, etc.) it maintained its own website, which at the time, in common with the websites of most UK museums, was the responsibility of its
marketing department, and was designed primarily as a means of attracting visitors to the museum building. Online information about the collections was rudimentary and in no way equated to an actual catalogue or database of holdings. However, a successful application to the New Opportunities Fund’s NOF-Digitise programme provided the Theatre Museum with the chance to use the web to make its collections accessible to a wider audience, and at the same time to develop expertise and an awareness of the demands of this new medium of communication.

The NOF-Digitise programme offered a total of £50 million for heritage organizations to digitize areas of their collections and make them available to new audiences. To qualify for funding, the websites created with this funding had to be geared towards young people, families, lifelong learners and those who were socially excluded due to their cultural or socio-economic background. Organizations such as museums, archives and libraries were identified as guardians of very rich collections but it was felt that there were barriers preventing wider use of these materials.

**PeoplePlayUK**

The Theatre Museum was well placed to respond to the call for applications. It held very large collections of highly visual material in a wide range of formats, spanning stage costume, photographs, designs, ephemera, ceramics, posters, puppets and more. Even better, much of this material was free of intellectual property rights (IPR) constraints, either by virtue of its age or because the rights were owned outright by the museum. It also had access to professional stage performers who formed the museum’s animation and interpretation team. Since we wished to exploit a ‘theatrical’ look and feel for the project and enliven the content by including audio recordings of music and spoken word specially commissioned for the site, this provided us with a valuable advantage.

The museum took the objectives and aims of the NOF-Digitise programme very seriously, carefully considering the target audiences and their needs. The website we devised, called PeoplePlayUK, would provide a combination of packaged content about our most popular subject areas (pantomime, musical theatre, drama, dance, music hall, opera, etc.) with a database of 1500 objects, each with an ‘object story’ (see, for example, Figure 3.1) to engage the visitor. Rather than merely providing a physical description of an
HUDSON: THE DIGITAL MUSEUM

object, this story would also try to draw on the ‘human interest’ associated with it. The souvenir booklet of Barnum & Bailey’s circus tour of America, 1897, is cut in the shape of an elephant. It features diary details of the day-to-day life of the circus on tour. Even the births of a leopard and of a sheep are recorded. The booklet gives a fascinating glimpse into the huge operation of mounting a tour. Fifty-three people are listed as selling tickets and acting as ushers. There are departments for railroad and baggage, men to erect the tents, special staff to care for the animals, others to maintain costumes and wigs. This souvenir may have been intended as an advert for the circus but also acts as a tribute to James A. Bailey who formed the circus partnership with P. T. Barnum in 1881.

For the specified audience, we felt that it was necessary to provide a rich mix of lively content, interpretation, glossaries, links and a multimedia approach to encourage visitors to engage with the content and explore the site.

Additional benefits

The project would not only provide an ideal means of showcasing the collections and of providing useful information of a high quality to its audience, it would help the museum to increase its efficiency by allowing visitors to answer their own enquiries rather than by using the museum’s overloaded e-mail and telephone enquiry services. After a two-stage application process that began in 1999 the museum was successful in being awarded £350,000, and the project was ready to begin in January 2002. The Theatre Museum’s PeoplePlayUK project was just one of 150 similar projects funded by NOF, some larger, some smaller. Most participants were novices, all starting to learn about digitization and website design at the same time. Although we had good technical advice from NOF’s technical advisory service, and within the V&A we could draw on highly expert IT and photographic departments, we had to design and manage the entire workflow, from the initial selection of objects through to quality control and user testing of the site, with no prior experience. Rather than using an external digitization bureau, all our scanning was done on our own premises (and still is) because moving museum objects off-site can be administratively complex. As an example of how basic it got, our scanners were selected from recommendations in What Scanner magazine.

NOF-Digitise was very much a public access project and not a retrospective cataloguing initiative. All projects had to employ Dublin Core-compliant metadata but at an early stage we realized that there would simply not be time to create records of the standard required by the V&A’s in-house cataloguing systems. We wished to use a mix of museum objects, library items, ephemera and archives to enliven the content and link very diverse materials. By rights, these materials would need to be catalogued in three different V&A systems. The added time and expense this would involve was prohibitive, so instead we created a dedicated database to contain the 1500 object records, their stories and their image metadata.

The PeoplePlayUK site was extremely popular with users – it achieved high visitor figures, averaging 170,000 visits per month at a time when the Theatre Museum’s own website was achieving barely a third of that figure. Although it was created for the non-
specialist it had many devotees among academics and experts because the quality of the information was very high, even though it was packaged in an accessible style.

Conclusion
By early 2009 security flaws were beginning to emerge that allowed hackers to launch malicious e-mails purporting to come from the V&A – a form of e-hijacking, and the V&A was told by the Joint Academic Network (JANET) to take the site offline because it posed a major risk. Since then, the V&A web team has been migrating the PeoplePlayUK content to the V&A website but minus some of the original site’s features and its separate branding.

The PeoplePlayUK project provided us with a very effective training experience that taught us all the basic skills needed for online publishing – workflow design, writing accessibly for the web, the creation and use of a style sheet to ensure consistent use of language, image manipulation, logical structuring of content and quality control. We also experimented with combining diverse materials in a complementary way and presenting them in a lively fashion, for example, by having an actor read a newspaper report rather than simply digitizing the press cutting.

The biggest and hardest lesson we learnt from this project was that it is essential that digital content resides within core institutional systems so that metadata can be updated in one single location, and content reused for other projects. It also ensures that it will be preserved alongside the organization’s other digital assets.

The enduring benefits of the NOF-Digitise £50 million total investment have been hotly debated. For the Theatre Museum, the investment was of undisputable value. Fifteen hundred of the Theatre Museum’s best objects were digitized to a high resolution and are still available for reuse from the V&A’s digital asset management system. Most of the museum objects digitized through this project have now been incorporated into the V&A’s online object catalogue, and the content written for the original website has been reused in publicly accessible web content. Not only was the content highly appreciated by its intended audience, it continues to be actively used, albeit through different channels of delivery.

The experience also vastly increased the Theatre Museum’s capacity for engaging with subsequent digital projects.

JISC-funded collaborative project: the East London theatre archive
Introduction
In 2006 the department was approached by the University of East London (UEL) and invited to participate in an application for a large-scale digitization project to be funded by the JISC. The focus of the project would be East London Theatre. The suggested partnership comprised our two organizations plus a number of theatrical venues and companies, including the Theatre Royal Stratford East, Hackney Empire, Hoxton Hall and the Half Moon Children’s Theatre. The technical partner was the Arts and
HUDSON: THE DIGITAL MUSEUM

Humanities Data Service (AHDS) Executive at King’s College London (which became the Centre for e-Research in 2008). The successful application resulted in a two-year digitization project that produced over 15,000 digital images (10,000 sourced from the V&A Theatre and Performance collections) plus a theatre gazetteer, interactive map and a series of introductory essays on themes such as animal performers, East End immigration and nautical drama.

Exploding the myths
The project began in March 2007 and the website (www.ELTA-project.org) was launched in March 2009. Of approximately £500,000 of project funding, the V&A received around £200,000, which allowed us to catalogue and digitize material that had been severely under-exploited in the past. Some of the 19th-century playbills we digitized had been inaccessible for years because of the fragility of the paper on which they were printed. In terms of resource discovery and access, this was therefore a very successful project. Traditionally, research interest has always focused on West End theatre. The East London theatre archive (ELTA) project exploded several myths concerning the nature of performance in East London (which was by no means confined to music hall and variety) and opened up new research opportunities for scholars in many disciplines including social, economic and urban history, as well as the more obvious disciplines of English literature and drama. This time we created standardized MARC catalogue records using the V&A’s bibliographical system, Horizon, from where they can be migrated when that system is ultimately superseded.

Conclusion
One of the more challenging aspects of this exercise, especially for the project managers at UEL, was the diversity of the partner organizations, all of whom were different in their size, status, aims and in what they could contribute to the project. Some were very small organizations with no member of staff designated to care for their archives. The partners’ degree of copyright awareness was also very uneven. UEL is to be congratulated for maintaining a productive partnership to the end of the project, despite the disparate aspirations and abilities of the participant organizations. Imposing deadlines for project milestones can be difficult if this essentially not-for-profit work conflicts with an organization’s commercial imperatives.

One interesting area of tension was the need to create an authoritative and scholarly resource for use by the research community while at the same time presenting the content in an accessible style. Primary sources such as playbills and account books do need a degree of interpretation to make them useful. The thematic essays certainly provided a way into new areas of research for those new to a subject area, and further work on using the site’s content for creating teaching and learning resources, and for evolving new works of performance inspired by the archive, is currently being carried out under a follow-up, JISC-funded initiative, called CEDAR (Clustering and Enhancing Digital Archives for Research, www.uel.ac.uk/lls/resources/cedar.htm). This initiative will encourage the
addition of user-generated content, and will present case studies illustrating the potential uses to which the digitized material can be put in order to create new resources.

V&A-funded digitization programme

Since 2007, the Department of Theatre and Performance has additionally received internal funding to digitize areas of its collections. Initially the intention was to balance the outputs between the mass digitization of largely two-dimensional copyright-free material, such as theatrical prints and photographic collections, and the creation of ‘value added’ learning packs in the style of PeoplePlayUK. The latter would provide web pages explaining and illustrating a specific aspect of performance history or practice. The first subject chosen was one we knew held great interest for audiences – that of designing, making and wearing theatrical costume. The project entailed selecting 50 of our most interesting costumes from a collection of around 3000, having them mounted on mannequins and photographed (converted in some cases to 360-degree animations), writing authoritative essays, captions and presenting the material in a logically structured package with links to related objects such as stage photographs and costume designs. Although we succeeded in completing the pack, the project was extremely time-consuming and was inevitably dependent on the intervention of the museum’s Web Team to code and publish it. The iterative process between web designer and client of presenting and signing off content worked well under the PeoplePlayUK model because it was fundamentally a commercial relationship defined by a formal contract. It was simply not appropriate when working in-house with a small and hard-pressed web development team, under constant pressure from competing internal projects such as web content relating to exhibitions, events and other time-specific activities.

By contrast, the department’s mass digitization projects, which have produced almost 50,000 digital images and associated catalogue records over the last three years, have been easier to accomplish. A simple workflow ensures that objects are photographed, catalogued to a minimum standard and published to the web without requiring input from the web design team. The launch of the V&A’s award-winning ‘Search the Collections’ initiative (http://collections.vam.ac.uk) has provided an effective and rapid system for connecting the public with descriptions and images of its objects.

National Video Archive of Performance recordings

One further area of digitization undertaken by the department has been that of the recordings made by V&A’s National Video Archive of Performance. Founded in 1992, the project has produced around 250 full-length recordings of major UK stage productions, recorded professionally to archive standard. Over this period, changes in video media have resulted in recordings being created in several different formats, for both long-term preservation and for viewing purposes. The digitization project will even out any inconsistencies and make the collection easier to manage and exploit in the future.

My organization has been fortunate in the range of digitization projects it has been able to undertake but our experiences have been far from unique. The following section will
set our findings within the broader context of the UK museums sector.

The bigger picture – digitization in UK museums

Over the past 15 years, UK museums have been engaged in digitizing their collections, with widely differing outcomes. In 1999, the National Museum Directors’ Conference published its landmark report, *A Netful of Jewels: New Museums in the Learning Age*. It foresaw a digital future for museums established on the foundations already being built. This would comprise a broad spectrum of digital initiatives, including:

- relevant participatory galleries and digital exhibits
- digital cameras, smart cards and other media for use during a visit
- content created by visitors as well as by museum staff
- facilities for searching the collections in ways that are relevant to visitors
- trained staff to help visitors learn
- interactive websites and online services
- online information to help in planning visits
- connections between the actual and the virtual museum, and with other cultural resources locally and worldwide.

The remaining section of this chapter will look at how far this vision has been fulfilled after a decade of significant investment in museum digitization initiatives.

Why digitize museum collections?

In analysing which areas of this vision have and have not been delivered and why, it may perhaps be useful first to consider what philosophical, operational and economic motives underpin museums’ efforts to digitize their collections. For example, the absolute requirement for museums to be accountable for all their objects, to know their exact location at all times, is not fully understood outside the sector. Digitization is a powerful tool in the auditing process and explains why museums may well seek to digitize their collections exhaustively, rather than focusing only on specific object types or themes.

An altruistic desire on the part of curators to make their collections widely accessible to audiences has to be the most obvious answer to this question of motivation. Although some museums reportedly still struggle with the conflict between making their collections accessible online and a feared reduction in real visitor numbers to their museums, the experience of those that have actively published their collections via the web is that this only builds their reputation and profile, and leads to larger numbers of visitors wishing to engage with the real objects and to participate in the events and activities designed around them.

Museums with a strong and well-populated web presence also enhance their reputations as centres of expertise. Most web visitors arrive at content through search engines, not through the home page, so a site that is rich in content optimized for search engines will achieve high visitor numbers. Larger museums, including the national