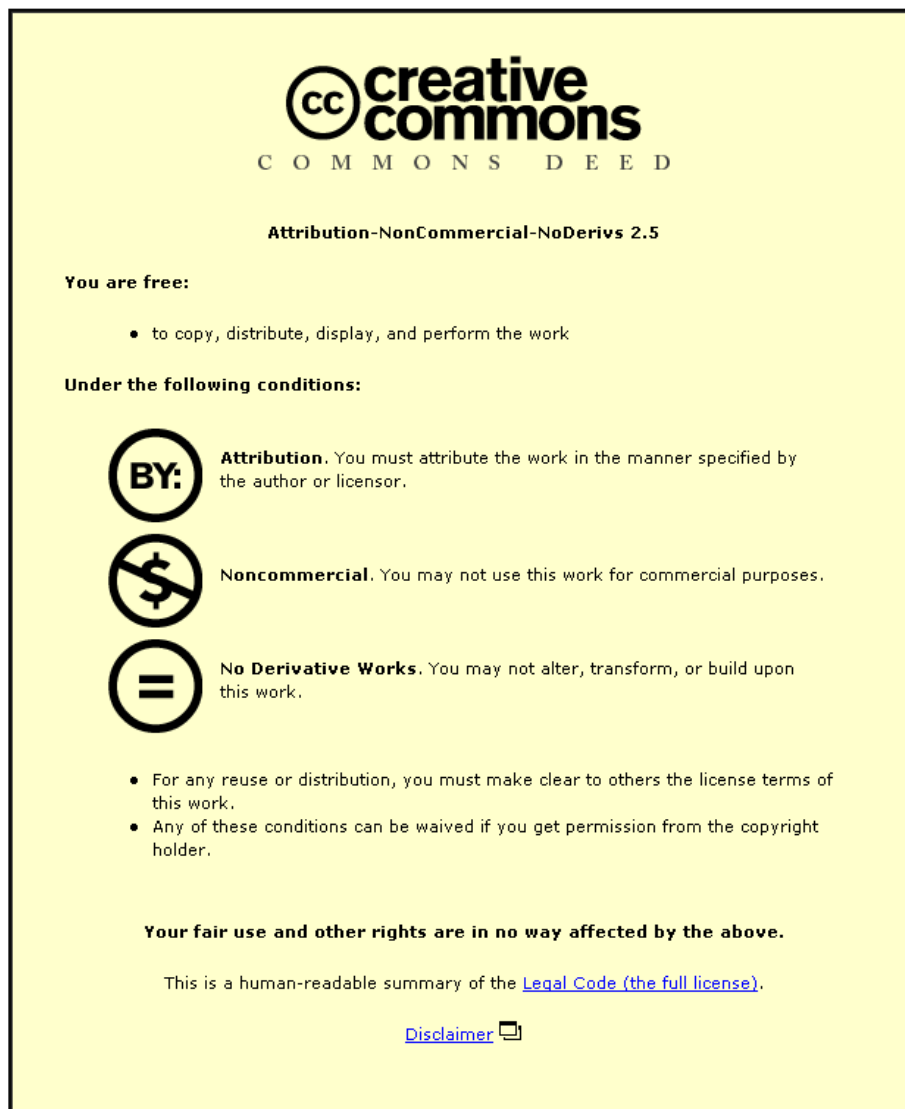


This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (<https://dspace.lboro.ac.uk/>) under the following Creative Commons Licence conditions.



For the full text of this licence, please go to:  
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

# Selection of digital material for preservation in libraries, archives and museums

by

J. Clare F. Ravenwood

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of  
Doctor of Philosophy of Loughborough University

May 2013

© J. Clare F. Ravenwood 2013

## ABSTRACT

---

Digital material has different preservation requirements than non-digital and is at greater risk of loss unless deliberate preservation activities are undertaken. Digital preservation is an on-going managed process designed to enable continued use of digital material for as long as necessary. Much of the digital preservation research to date has focused on understanding technical steps in preserving digital objects and there has been less attention paid to assumptions about selection and the conceptual underpinnings of practice. Selection is done, not conceptualised. Therefore the aim of this research was to investigate the theory and practice of selection for digital preservation in UK memory institutions. The objectives employed to achieve this aim were firstly to examine the underlying theory relating to selection in libraries, archives and museums of non-digital material. The research then went on to investigate who the stakeholders are in selection, how selection of digital material is performed and identifying the key influential factors in selection. An intensive, qualitative approach was used to complete these objectives. A thorough review of the literature provided a theoretical background to selection in libraries, archives and museums. Then preliminary data were gathered through a set of exploratory interviews with eight digital preservation 'experts' in order to provide an overview of selection for digital preservation. The findings from these interviews then formed the basis for the second set of interviews with twenty five practitioners working in libraries, archives and museums. The views of practitioners were under-explored in the literature although it is they that perform selection. In addition to these interviews, twenty two current digital preservation policies were examined.

This research has found that there is on the whole little change required for selecting digital material, in comparison to selecting non-digital material, although technical criteria relating to the ability of the institution to manage and preserve the material are of high importance. There is a clear assumption in institutions of selection leading only to permanent collecting, which should be questioned. This research has uncovered drivers to selection, including external funders, and barriers, which include a lack of confidence and knowledge on the part of practitioners in how to select and manage digital material. Concepts identified through this research provide a deeper understanding of selection for digital preservation in different contexts and encapsulate key factors underpinning selection. The concept of professionalism is a key factor; the need to be professional and ethical guide's practitioners through specific professional skills and knowledge. The practitioners become engaged with digital material and the level of engagement mirrors the way digital material is conceptualised by practitioners. Many stakeholders were identified, including managers, senior managers, users, creators and donors, funders, other organisations and IT staff. Relationships with stakeholders and the possible roles they play in selection were found to be key factors in selection. These findings contributed to the achievement of the final objective, which was to develop a conceptual model

of key factors underpinning selection for of digital material for preservation. The conceptual model consists of five main concepts and their relationships: professionalism; relationships; organisational capabilities; material properties; and boundaries. There is a clear need for greater availability and access to training and networking opportunities for practitioners in order to increase engagement with digital material.

Through this research, factors relating to selection have been identified and conceptualised. It has uncovered issues not previously addressed, in particular relating to the social aspect of selection. This research provides an understanding of the complexities of selection and the influences upon it.

Key words: selection, criteria, digital preservation, digital material, libraries, archives, museums.

# CONTENTS

---

Abstract .....	i
Contents .....	iii
Tables and Figures.....	viii
Acknowledgements .....	ix
Glossary and Abbreviations .....	x
Chapter 1 Introduction.....	1
1.1 Cultural heritage .....	2
1.1.1 The cultural heritage sector in the UK.....	3
1.2 Digital material.....	5
1.3 Digital preservation.....	7
1.4 Selection.....	8
1.5 Research rationale .....	9
1.6 Scope .....	10
1.7 Aim and objectives .....	11
1.8 Thesis structure.....	12
Chapter 2 Literature review .....	14
2.1 Selection in memory institutions.....	15
2.1.1 Archival appraisal .....	15
2.1.2 Selection in libraries .....	18
2.1.3 Selection in museums .....	22
2.1.4 Summary .....	24
2.2 Digital preservation.....	26
2.2.1 Technical aspects.....	27
2.2.2 Significant properties.....	30
2.2.3 Summary .....	33
2.3 Selection of digital material for preservation .....	33
2.3.1 Criteria .....	34

2.3.2	Value .....	38
2.3.3	Roles and responsibilities.....	41
2.3.4	Finance and costs .....	46
2.3.5	Legal issues, ethics and risk .....	49
2.3.6	Policy .....	51
2.3.7	Summary .....	56
2.4	Conclusions .....	57
Chapter 3 Research methodology .....		60
3.1	Research philosophy .....	61
3.2	Research design .....	64
3.2.1	Qualitative research .....	64
3.2.2	Data collection .....	68
3.2.3	Interviews .....	70
3.2.4	Quality in qualitative research .....	72
3.2.5	Ethical considerations .....	73
3.3	Literature review.....	75
3.4	Expert interviews .....	76
3.4.1	Sampling.....	76
3.4.2	Interview design .....	76
3.4.3	Data analysis .....	77
3.5	Practitioner interviews .....	79
3.5.1	Sampling.....	79
3.5.2	Interview design .....	83
3.5.3	Data analysis .....	83
3.6	Policy analysis.....	83
3.6.1	Sampling.....	84
3.6.2	Data analysis .....	85
3.7	Summary .....	86
Chapter 4 Expert findings.....		87
4.1	Professional roles and responsibilities .....	88

4.1.1	Ethics.....	90
4.1.2	Knowledge and skills.....	91
4.2	Institutional issues.....	92
4.2.1	Roles and responsibilities.....	92
4.2.2	Risks.....	96
4.2.3	Legal issues.....	99
4.2.4	Policies.....	100
4.2.5	Finance and costs.....	101
4.3	Conceptual issues.....	102
4.3.1	Significant properties.....	102
4.3.2	Value and criteria.....	103
4.3.3	The future.....	106
4.4	Conclusions.....	107
Chapter 5 Practitioner interview findings.....		110
5.1	Roles and responsibilities of the professional.....	111
5.1.1	Selection as a professional role.....	111
5.1.2	Determining value.....	111
5.1.3	Subjectivity.....	114
5.1.4	Legal and ethical behaviour.....	115
5.1.5	Knowledge and skills.....	117
5.1.6	Engagement.....	120
5.2	Stakeholder roles, responsibilities and relationships.....	121
5.2.1	Sources.....	123
5.2.2	Guides and limiters.....	126
5.2.3	Users.....	129
5.2.4	Collaborators.....	130
5.2.5	Stakeholders and roles.....	131
5.3	Reactions to digital.....	132
5.3.1	Technical challenges.....	135
5.3.2	Significant properties.....	136

5.4	The selection process .....	138
5.4.1	Drivers to selection .....	141
5.4.2	Criteria .....	143
5.4.3	Outcomes .....	145
5.5	Institutional factors .....	146
5.5.1	Mission .....	146
5.5.2	Policies in practice .....	147
5.5.3	Policy .....	149
5.5.4	Legal, ownership and risk .....	151
5.5.5	Finance and resources .....	153
5.6	Conclusions .....	156
Chapter 6 Discussion .....		161
6.1	Conceptual model of key factors .....	162
6.1.1	Boundaries .....	164
6.1.2	Capability and properties .....	166
6.1.3	Professionalism .....	166
6.1.4	Relationships .....	169
6.2	Stakeholders .....	170
6.3	Practice .....	176
6.3.1	Criteria .....	177
6.3.2	Drivers and barriers .....	179
6.3.3	Policies .....	180
6.3.4	Assumptions and change .....	181
6.4	Summary of key findings .....	183
Chapter 7 Conclusions .....		187
7.1	Research objectives and main findings .....	188
7.1.1	Existing theory .....	188
7.1.2	Stakeholders .....	190
7.1.3	Practice .....	192
7.1.4	Key factors .....	193

7.1.5	Conceptual model of key factors .....	193
7.2	Recommendations .....	194
7.3	Research contribution .....	195
7.4	Suggestions for further work .....	195
7.5	Reflection and limitations .....	196
7.6	Concluding remarks .....	197
	Bibliography.....	198
	Appendices.....	230
	Appendix 1 .....	230
	Appendix 2.....	231
	Appendix 3.....	232
	Appendix 4.....	233
	Appendix 5.....	234
	Appendix 6.....	235

## TABLES AND FIGURES

---

Table 1 Criteria.....	37
Table 2 Research paradigms .....	62
Table 3 Intensive and extensive research .....	65
Table 4 Qualitative methodologies .....	66
Table 5 Interviews .....	71
Table 6 Telephone and email interviews.....	72
Table 7 Positivist quality criteria .....	72
Table 8 Techniques to ensure quality .....	74
Table 9 Search terms .....	75
Table 10 Practitioner interview numbers.....	80
Table 11 Practitioner interviews by setting and role .....	81
Table 12 Practitioner interviewees .....	82
Table 13 Initial practitioner coding framework .....	83
Table 14 Sources .....	124
Table 15 Stakeholders and roles .....	131
Table 16 Drivers .....	141
Table 17 Key factors .....	162
Table 18 About the policy guides .....	231
Table 19 Comparison of policy guides .....	232
Table 20 Practitioner interview schedule .....	233
Table 21 Stakeholders .....	234
Table 22 Comparison of policies and policy guides .....	235
Figure 1 Engagement.....	120
Figure 2 Selection points.....	140
Figure 3 Model of key factors.....	163
Figure 4 Responsibilities .....	174

## **ACKNOWLEDGEMENTS**

---

I am grateful to the Department of Information Science at Loughborough University for the opportunity and funding to undertake this research. Thanks must go to my ever patient supervisors, Professor Graham Matthews and Dr Adrienne Muir, who must have wondered what they had got themselves into. They have encouraged me, guided me and cajoled me into achieving more than I thought I could. I am also grateful to the participants in this research, especially the practitioners, who were generous with their time and passionate about their work.

On a personal note, this is for my parents; one of whom taught me to read and love books before I went to school and the other who was proud of my academic accomplishments. Finally I have to thank Simon, without whose love and encouragement I would not have started on this road in the first place.

## GLOSSARY AND ABBREVIATIONS

---

ACE	Arts Council England
ADS	Archaeology Data Service
ALA	American Library Association
APARSEN	Alliance for Permanent Access to the Records of Science Network
ARA	Archives and Records Association
BBC	British Broadcasting Corporation
BL	British Library
BnF	Bibliothèque nationale de France
BRTF	Blue Ribbon Task Force
CARARE	Connecting Archaeology and Architecture in Europeana
CCSDC	Consultative Committee for Space Data Systems
CEDARS	CURL exemplars in Digital Archives
CHIN	Canadian Heritage Information Network
CILIP	Chartered Institute of Library and Information Professionals
CLIR	Council for Library and Information Resources
CURL	Consortium of University Research Libraries
CyMAL	Museums, Archives and Libraries Wales
DCC	Digital Curation Centre
DCMS	Department for Culture Media and Sport
DIFFUSE	Dissemination of InFormal and Formal Useful Specifications and Experiences
DISCmap	Digitisation in Special Collections: mapping, assessment, prioritisation
DPC	Digital Preservation Coalition
DPE	Digital Preservation Europe
DRAMBORA	Digital Repository Audit Method Based On Risk Assessment
DTPT	Digital Preservation Training Programme
EROS	Earth Resources Observation and Science
ERPANET	Electronic Resource Preservation and Access Network

EU	European Union
HLF	Heritage Lottery Fund
ICA	International Council on Archives
ICOM	International Council of Museums
ICPSR	Interuniversity Consortium for Political and Social Research
IFLA	International Federation of Library Associations and Institutions
InSPECT	Investigating Significant Properties of Electronic Content
INTERPARES	International Research on Permanent Authentic Records in Electronic Systems
IPO	Intellectual Property Office
IT	Information Technology
JISC	Joint Information Systems Committee
KEEP	Keeping Emulation Environments Portable
KRDS	Keeping Research Data Safe
LAC	Libraries Archives Canada
LDAP	Legal Deposit Advisory Panel
LIFE	Life Cycle Information for E-Literature
LISA	Library and Information Science Abstracts
LOCKSS	Lots of Copies Keep Stuff Safe
MA	Museums Association
MLA	Museums and Libraries Association
NINCH	National Initiative for a Networked Cultural Heritage
NLNZ/ANZ	National Library of New Zealand/Archives New Zealand
OAIS	Open Archival Information System
openDOAR	Directory of Open Access Repositories
PADI	Preserving Access to Digital Information
PANDORA	Preserving and Accessing Networked Documentary Resources of Australia
PARADIGM	Personal Archives Accessible in Digital Media
PARSE.Insight	Permanent Access to the Records of Science in Europe
Pdf	Portable document format

PLANETS	Preservation and Long-term Access through Networked Services
POWR	Preservation of Web Resources
PREMIS	Preservation Metadata: Implementation Strategies
PRONI	Public Record Office of Northern Ireland
PRONOM	a web-based technical registry to support digital preservation services developed by TNA
RCAHMS	Royal Commission on the Ancient and Historical Monuments of Scotland
REM	Renaissance East Midlands
RI	Representation Information
RLUK	Research Libraries UK
SPECTRUM	The UK Museum Collections Management Standard
TNA	The National Archives
TIMBUS	Timeless Business Processes and Services
TRAC	Trusted Repositories Audit and Certification
UKOLN	a centre which advises on digital infrastructure, information policy and data management, based at the University of Bath, part funded by JISC
ULCC	University of London Computer Centre
UNESCO	United Nations Education Scientific and Cultural Organisation

## CHAPTER 1 INTRODUCTION

---

This chapter presents the background to the research, defines key terms, describes its scope and significance, and presents the aims and objectives. The final section describes the structure of the rest of the thesis.

## 1.1 Cultural heritage

Future cultural heritage is at least in part dependent on the selection decisions made now within memory institutions traditionally tasked with preserving it. The selection for preservation process creates heritage, as criteria determine what is to be valued and defined as heritage (Cameron 2008, p.177). Selection of digital material for preservation transforms it into 'heritage', something to be passed to the future, and so it is worthy of the large amount of resources that are needed to keep it continually useable into the future, providing justification for the deployment of those resources. Heritage is something that is created rather than merely possessed (Smith 2006, p.304). Heritage is made, not found, and choices made by those selecting material for preservation determine what will be considered valuable enough to be heritage. Material is heritage because it is selected, not only selected because it is heritage.

That heritage is a political and cultural process is widely acknowledged in the literature (for example Graham and Howard 2008, p.2; Hall 2005, p.24; Silberman 2008, p.82). Heritage involves passing on knowledge to future generations. As Gorman (2007, p.286) states

*The term "cultural heritage" contains within it a clear implication – that of onward transmission. The word "heritage" means something transmitted by or acquired from a predecessor. In order for that generational transfer to take place, the item of cultural heritage must be recorded and preserved.*

'Cultural heritage' is a broad term. Holden and Hewison (2004, p.12) describe how the National Heritage Act (1980), subsequent legislation, policy directions and successive trustees of the Heritage Lottery Fund (HLF) have all deliberately not given a definition of 'heritage', preferring to leave the definition implicit in the activities it is involved with and the applications it receives. The HLF takes a broad view of what is heritage, as is reflected in the guidance given on the HLF website. In this (HLF 2013a) heritage includes:

- People's memories and experiences (often recorded as 'oral history')
- Histories of people, communities, places and events
- Languages and dialects
- Cultural traditions
- Historic buildings and townscapes
- Archaeological sites
- Collections of items, archives or other materials
- Natural and designed landscapes habitats and species
- Sites and collections linked to our industrial, maritime (to do with the sea) and transport history.

Heritage includes collections of materials found in libraries, archives and museums. The term 'memory institutions' or 'cultural heritage institutions' may be used to describe these, though the latter term has the problem of including a wider range of organisations. Within the cultural heritage sector there seems to be no agreed collective term; the acronym 'GLAMs' may be used, referring to 'Galleries, Libraries, Archives and Museums', or merely 'LAMs' - Libraries, Archives and Museums. A further collective term that may be used is 'memory institution'. A useful definition of a memory institution is from Dempsey (1999)

*Archives, libraries and museums are memory institutions: ...Their collections contain the memory of peoples, communities, institutions and individuals, the scientific and cultural heritage, and the products throughout time of our imagination, craft and learning. They join us to our ancestors and are our legacy to future generations.*

This definition is useful as it links archives, museums and libraries to culture, heritage and cultural memory. Thus it is used in this research, which focuses on digital preservation activities in libraries, archives and museums i.e. memory institutions. Whilst the issues relating to digital preservation are not confined to cultural heritage artefacts but are applicable to a wider range of material, it is often libraries, archives and museums that have a responsibility for selecting, collecting and preserving it. It should be noted that although document based heritage is mostly associated with libraries and archives, museums are increasingly becoming engaged with material in digital form and are included in this research for this reason.

### **1.1.1 The cultural heritage sector in the UK**

Each country that makes up the UK has its own heritage governance structure and organisations. In England, the Department for Culture, Media and Sport (DCMS) has responsibility for the arts, cultural property, libraries, museums, galleries and the historic environment, along with other areas such as gambling, sport and tourism (DCMS, n.d.a). The DCMS funds agencies with specific responsibilities in the cultural sector. English Heritage is an 'Executive Non-departmental Public Body', sponsored by DCMS, which acts as the government's statutory advisor for the historic environment (DCMS n.d.b) and has a responsibility to care for and protect the historic environment. The Arts Council England (ACE n.d) has responsibility for supporting and developing the arts in England, which includes libraries and museums following the demise of the Museums Libraries and Archives Council (MLA) in 2011. The ACE also administers the 'Designated Collection' scheme, whereby nationally important collections held in non-national institutions are recognised and awarded 'designated' status. Originally this applied only to museum collections, but was expanded to include archives and library collections in 2005. Currently there are one hundred and forty designated collections, such as those at Birmingham Central Library, Derby Museum and Art Gallery and Norfolk Record Office. In addition ACE took responsibility for administering the

Museum Accreditation Scheme from the MLA, in partnership with CyMAL (Museums, Archives and Libraries Wales), Museums Galleries Scotland and the Northern Ireland Museum Council.

In Scotland responsibility for cultural heritage lies with the Cabinet Secretary for Culture and External Affairs who has responsibility for culture and the arts, the built heritage and Historic Scotland as well as Europe and external affairs (The Scottish Parliament n.d.). The Directorate for Culture, External Affairs and Tourism acts as an adviser to Ministers and presents and delivers policy. The Scottish government funds other bodies that have a stake in the cultural heritage sector, including the Scottish Arts Council, Historic Scotland and the Royal Commission on the Ancient and Historical Monuments of Scotland (RCAHMS). Government funded national institutions include the National Archives of Scotland, The National Library of Scotland, National Museums Scotland and The National Galleries of Scotland (The Scottish Government 2013). In Wales responsibility for museums libraries and archives lies with the Minister for Heritage (Welsh Government n.d.a). CyMAL is part of the Welsh Government that provides the Minister with policy advice and supports museums, libraries and archives in Wales (Welsh Government n.d.b). In Northern Ireland the responsibility for the arts, museums, public archives and libraries is with the Department of Culture, Arts and Leisure, part of the Northern Ireland Executive (Department of Culture, Arts and Leisure n.d.). This department has oversight of non-departmental public bodies including the Northern Ireland Museums Council, Public Record Office of Northern Ireland (PRONI) and Libraries NI (Department of Culture, Arts and Leisure. n.d.). There are further important contributors to heritage in the UK including the Heritage Lottery Fund, which is a non-departmental government body that distributes part of the funds raised through the National Lottery for Good Causes (HLF 2013b). There are also a large number of charities in the UK which have cultural heritage in their remit, focusing particularly on the built heritage, such as the Pilgrim's Trust, the National Trust and the National Trust for Scotland.

The public archive sector comprises approximately two thousand archive services (TNA n.d.)The National Archives (TNA) is the government's national archive for England, Wales and the United Kingdom.TNA is a government department and an executive agency of the Ministry of Justice. It acts as the central advisory body for records and archives, manages Crown copyright, and undertakes other related activities (TNA n.d.).Since October 2011 it has taken on responsibility for archives from the now defunct MLA. Many archives are held not only by local authorities, but also corporations, universities, religious organisations, and a wide array of private organisations, societies, charities and individuals.

Policy responsibility for public libraries lies with the DCMS but delivery is the statutory responsibility of the one hundred and fifty one local authorities. The British Library is an executive non-departmental public body of the DCMS (DCMS n.d.c). There are approximately three and a half thousand public libraries in the UK (Smithies 2011, p.5). There are also

libraries in a wide variety of settings, including universities, schools, further education colleges, the NHS, in businesses or law firms, government libraries, and there are some independent subscription libraries which are privately owned but allow public use for a fee.

The DCMS sponsors twenty-one museums and galleries of which many are national museums such as the British Museum, the Imperial War Museum and the Royal Armouries (Gov.uk, n.d.) Figures available from the Arts Council's 'A review of research and literature on museums and libraries' (Smithies 2011, p.4) reveal that there are about sixteen hundred museums in the UK, of which fourteen hundred are accredited. Local authorities fund about a third of local museums and there is a wide independent sector funded by charities and other third sector organisations.

## 1.2 Digital material

Digital material has particular properties and requirements, different to those of traditional non-digital based material, which drive the need for a different approach to preservation. Digital material is machine and software dependent thus the data object alone does not preserve the information, but the means to process the data into meaningful information must also be preserved. Digital material depends on changing technology, where both hardware and software can become obsolete. Digital media is fragile; magnetic devices are vulnerable to a range of environmental factors as are optical storage devices and hard drives, which can also fail mechanically. There is an enormous amount of digital material created every year; the most recent annual 'Digital Universe' report available (Gantz and Reinsel 2011) estimates that in 2011 1.8 zettabytes, equivalent to 1.8 trillion gigabytes, of digital data will be created or replicated. The form of digital heritage material may vary widely. As Owen (2007, p.46) states '*digital heritage materials are not just traditional heritage materials in digital form*'. Digital material is global, it may originate from anywhere and geographic boundaries may not apply (Owen 2007, p.46). Digital material is dynamic and interactive; it can be constituted through social interaction and is linked in multiple ways through networks. These factors mean that digital material is complex to preserve and keep useable over time, so traditional methods and approaches to preservation may no longer be appropriate, especially in the long term. Strategies to preserve digital material have been a particular focus of research (see section 2.2).

A general definition of digital heritage is as '*a selected pool of material in a digital format deemed worthy of preservation for posterity*' (Cameron 2008, p.172). Whilst this definition links digital material with concepts of heritage and value, it is also useful to turn to the digital preservation literature for more specific definitions. Thibodeau (2002) describes the digital object as having three essential properties, all of which have to be addressed when considering preservation. The first aspect of a digital object is simply '*an inscription of signs on a medium*', which have to be managed by a storage system and have no meaning without another level of processing. The digital object is also a logical object, which is a '*unit recognised by some*

*application software*'. This describes the way the object is processed by software to make it usable, usually by reference to the type of data it is such as a JPEG image, or a Word document. Finally there is the level of the 'conceptual object', which is the item we recognise as 'a *meaningful unit of information*' such as a document or photograph. In order to render a digital object we must be able to identify and process the physical and logical components to reconstitute the conceptual object effectively.

Though this research is concerned mainly with the preservation of digital material, it is useful to consider the practice of digitisation. Put simply, digitisation is the process by which an analogue item is converted into a digital format (Youngs 2001, p.4). This produces a 'surrogate' copy which can be used in lieu of the original; this has been previously carried out in libraries using other media such as microfilm (Edwards, Matthews and Nankivell 2000, p.2). There is a distinction in the literature between preserving 'digital heritage' and 'digitised heritage'. The former usually refers to the preservation of born digital material that has no analogue equivalent, whilst the latter refers to preserving analogue material by digitising it, such as a scan of an old manuscript or a 3-D rendition of an archaeological landscape. Conway (2010, p.64-65) explains the difference between digital preservation and preserving material by digitising it:

*Digitization for preservation creates valuable new digital products, whereas digital preservation protects the value of those products, regardless of whether the original source is a tangible artefact or data that were born and live digitally. Digitization for preservation and digital preservation are intimately related, but the underlying standards, processes, technologies, costs, and organizational challenges are quite distinct.*

Digitisation is used to aid access to fragile or valuable analogue material. Many surveys that have been performed on the state of preservation readiness and on preservation activities (for example Simpson 2005; Waller and Sharpe 2006) demonstrate that though many institutions engage with digital material, this may be in the form of or in response to short term digitisation initiatives to improve access. In considering digital preservation, this research considers digitised material and born digital material as very similar as both may need preserving. The definition of digital materials found in the Digital Preservation Coalition (DPC) Handbook (2008, p.24) is appropriate as it encompasses different categories of material created in different ways in one definition:

*A broad term encompassing digital surrogates created as a result of converting analogue materials to digital form (digitisation), and "born digital" for which there has never been and is never intended to be an analogue equivalent, and digital records.*

### 1.3 Digital preservation

Preservation can be considered a part of the wider process of curating digital material. According to Feeney (1999, p.5) the development of digital preservation as a focus for active interest originated from a Commission on Preservation and Access and the Research Libraries Group report from the USA (Waters and Garrett 1996). This report was analysed by Matthews, Poulter and Blagg (1997) for relevance to the UK context and they produced recommendations for action. Furthermore, a conference was held at Warwick by the JISC and the British Library (BL) which also produced an influential report (Fresco 1996). Similarly Beagrie (2006) traces the development of the term 'curation', ascribing it to a seminar in 2001 where the term was carefully chosen '*to explicitly transfer existing curatorial approaches to digital collections*' from libraries and museums (2006, p.5). The current definition of 'digital curation' given by the Digital Curation Centre (DCC 2010) is '*Digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle.*' The DCC have presented a 'lifecycle model' of curation (Higgins 2008), which includes preservation in two ways. Firstly as 'preservation actions' constituting a step in the lifecycle, which are specific actions needed to preserve material at a point in time, for example by refreshing media (see section 2.1.1). Secondly preservation is used in combination with curation as an on-going process meaning to '*Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle.*' (Higgins 2008, p.137). This statement is vague; the dual use of the term preservation in the model may cause confusion although it reflects the wider digital preservation literature in which 'preservation' may have more than one meaning - an on-going activity or specific actions. This is unhelpful; preservation should be viewed as an on-going activity:

*Digital curation and data preservation are on-going processes, requiring considerable thought and the investment of adequate time and resources.* (Digital Curation Centre (DCC) 2010)

Preservation is a process, not a one off action. A definition of digital preservation that is often cited is from the Digital Preservation Coalition (DPC). Jones and Beagrie (2001) produced a handbook, now updated online by the DPC, which broadly defines digital preservation as '*the series of managed activities necessary to ensure continued access to digital materials for as long as necessary*' (DPC 2008, p.24). This definition is clear and to the point, and it recognises preservation as an on-going activity rather than a one-off action. Another useful, and short, definition of preservation is from the American Library Association (ALA)(2009) '*Digital preservation combines policies, strategies and actions that ensure access to digital content over time*'. The definition of preservation used here is quite broad and is not limited to any particular preservation action or intervention; it highlights that digital preservation should include policies and strategies in combination with actions to be successful. This distinction between preservation which refers to the continuous process of preservation including planning and

management and 'preservation actions' which refer to specific actions is used throughout this research.

## 1.4 Selection

According to the Oxford English Dictionary, 'to select' means '*to choose or pick out in preference to another or others*'; this implies a judgement is made on the relative value of one thing to another. A reason for selection suggested in the literature is that of quality control (Harvey 2005, p.55). This reflects the view of libraries that one should only collect the best material, that which is of a high quality. This has been a long held view of librarians; Anderson (1999, p.76) commented

*The process [of selection] must be built and nurtured by knowledgeable individuals, mostly subject specialists who will ensure the intellectual integrity and worthiness of a digitized collection. Fortunately, librarians have been rendering this service for decades using traditional formats.*

Obviously this aspect can steer a dangerous course as it immediately asks by what standards is quality measured and who decides on the standards?

A difficulty with someone making judgements is that criteria, significance and value are culturally and subjectively bound. What is selected will reflect what the selector thinks is valuable and future digital heritage will reflect the selector's choices. This issue has been addressed in the wider heritage literature. Smith (2006, p.29) argues from a critical perspective that heritage, in particular material heritage, is about power, as it is the territory of experts, such as historians or archaeologists, who have authority and knowledge that others do not. This view is supported by Throsby (1997, p.14), who argues that in the definitions given of cultural heritage, for example by UNESCO, the determination of what constitutes 'culturally significant' or 'outstanding' relies on experts specifying the meaning of the values underlying these terms. Similarly Gibson (2009, p.75-76) demonstrates from a built heritage perspective that the language used in heritage documents, regardless of the well intentioned articulation of inclusive and consultative principles, undermines the commitment to these principles by putting communication with other stakeholders second to the evaluation and determination of significance by experts. Significance is not only culturally bound but can only be assessed by those who have specific knowledge or skills. However with the advent of digital technology traditional forms of cultural object creation are changing and the digital cultural heritage is much more complex due to the ability of non-expert communities to create their own material. Gracy (2007) examines the concept of moving image archives created by users, lying outside institutions, as are being seen in the Internet environment with YouTube or the Internet Archive for example (2007, p.193). She acknowledges that there are difficulties with this model, especially with copyright legislation leading to the need to control content and economic drivers

toward control by corporations (2007, p.194-195), but she asks the important and more general question of whether libraries and archives are still in possession of curatorial authority over cultural heritage, or if they should recognise the role others have in defining what is valuable enough to preserve?

## 1.5 Research rationale

This work was in part prompted by the researcher's interest in selection and partly through the question raised during previous reading of the digital preservation literature of '*who does the selecting?*' Perhaps understandably much of the digital preservation research has focused on understanding technical steps in preserving digital objects. Whilst these aspects are now much better understood, to date there has been less attention paid in the literature to assumptions about the practice of selection and the conceptual underpinnings of practice. The evaluation of digital material for preservation purposes seems in particular to not have been explored in a theoretical or conceptual way. Selection is performed, not conceptualised. When selection is addressed it is in terms of suggesting practical advice, such as appropriate criteria, and few commentators have addressed changes that may be needed to theory, policy and practice. Little research examining the criteria used for preservation has been undertaken and this research will consider the types of value and criteria appropriate to digital material. Consideration of concepts, such as 'value', in relation to the process of selection leads directly to further questions, such as where responsibility for evaluation in a digital networked environment lies. The views of practitioners on the question of selection have been under explored; their expectations and experiences of change from the introduction of digital material into their institutions have not yet been considered, though they are the people performing selection activities. Selection theory and practice in different types of memory institutions (referred to in this thesis as different 'domains') was developed over time for non-digital material and although the applicability of traditional selection criteria have been questioned, little has been done to investigate further.

Selection is seen as expensive (BRTF 2010 p.46; Lunghi et al 2012, p.218), yet necessary, although it is unclear whether traditional processes or strategies are effective in selecting digital material. The issue of selection is relevant to many current questions, underpinning issues such as:

- the best use of limited resources within institutions to focus on the most valuable material
- the development of curricula for digital preservation training to develop practitioner (i.e. those staff working in libraries, archives or museums who have selection as part of their job role) skills and knowledge
- the management of collections and their development

When the increasing amounts of digital material available is considered it seems timely to examine the underlying concepts and assumptions of the theory and practice of selection.

## 1.6 Scope

It was decided at an early stage that this research would focus on digital preservation in the UK, in order to make it manageable. However this would be explored within the wider international context of digital preservation, learning from and drawing on perspectives, research and practice from other countries. This is especially relevant considering the networked environment with which much digital preservation is concerned. The emphasis is on preservation, but theory and practice relating to digitisation will also be considered.

The research focuses on the cultural heritage sector, but again will be informed by work from a broader preservation context. There are many stakeholders with influence on and interest in preservation at different stages (Beagrie and Greenstein 1998). This research could not include all of them due to time and resource constraints, so the focus is on the institution. This is appropriate as it is memory institutions - libraries, archives and museums - which are tasked with preserving material for the future (Feather 1996, p.58; Usherwood, Wilson and Bryman 2005). Careful consideration was given to which stakeholders within the institutions were appropriate to include and it was decided to focus on practitioners and managers who have a close interest in selection as they have been rarely consulted before for their views. This has kept the research within manageable limits given the time and resources available.

The types of digital material referred to in this research are those produced from digitising paper-based items and 'born digital' material which has no non-digital equivalent. Whilst different forms of intellectual heritage material may have different preservation needs, especially audio-visual material (Wilson et al 2006, p.4-5), it is beyond the scope of the present research to address the research objectives through a detailed consideration of specific forms, and so they will be considered similarly as 'digital material'. The types of digital material relevant to this research include (though are not limited to):

- Documents
- Research data, databases, data sets
- Journals and e-books
- Government data
- Websites
- Social networking, blogs and email
- Sound and moving images
- Photographs
- Software

## 1.7 Aim and objectives

This research has the following overall aim: to investigate the theory and practice of selection for digital preservation in UK memory institutions. The aim is deliberately broad because of the exploratory nature of the research. It aims to analyse theories, policies and practice within memory institutions to identify stakeholders and key factors in selection to reflect the particular requirements of digital material. This research will add to the understanding of selection, detailing who is involved, how selection is carried out, what concepts are important, and what underlying factors influence selection. The following research objectives, with relevant research questions, will contribute to the achievement of the overall aim:

**Objective 1:** Achieve an overview and understanding of the theory of selection for preservation within UK memory institutions.

- What is the context in which selection takes place?
- What are the different concepts involved in selection?

In order to understand change and the impact of digital material on both theory and practice it is first necessary to gain an overview of the theory and practice relating to traditional non-digital material. This objective specifically looks at the theory behind selection practice in order to provide the theoretical context for objective 3.

**Objective 2:** Identify internal and external stakeholders in selection for digital preservation in memory institutions.

- Who has an interest in or influence on selection for preservation?
- What are their roles and responsibilities, both in theory and practice?

Traditional selection is carried out by people working in institutions, according to particular laws, policy and procedures, so it is important to gain an overview of these practitioners. But practitioners do not work in isolation and in order to understand the wider context of selection other stakeholders should be identified and their roles and responsibilities investigated.

**Objective 3:** Investigate the practice of selection for digital preservation in different UK memory institutions through the examination of practitioner views.

- How is digital material selected for preservation in those institutions which have a digital preservation function?
- What drivers and barriers are there to selection in different memory institutions?
- How do selection policies relate to selection practice?
- What assumptions have been made by memory institutions in their current thinking about selection?

- What changes to practice do practitioners perceive are necessary to select digital material for preservation?

This objective increases understanding of the practice of selection in institutions, considering digital material from the point of view of those who are carrying out selection.

**Objective 4:** Identify and describe the key factors in selection for digital preservation in memory institutions.

- What practical and theoretical factors influence selection practice for digital material?

In order to further the conceptualisation of selection the factors which influence it need to be identified and understood.

**Objective 5:** Construct a conceptual model of key factors and their relationships which influence selection.

In summary this thesis focuses on exploring selection for digital preservation, identifying the stakeholders, and examining theory, practice and policies of selection. Factors that influence policy and practice are explored from the perspective of both digital preservation experts and practitioners in the cultural heritage sector. A greater understanding of the similarities and differences between selection of traditional and digital material will allow stakeholders to manage the practice and policies of selection more effectively. By constructing a conceptual model of key factors this research will provide a theoretical underpinning for selection.

## 1.8 Thesis structure

Chapter one has described the aims and objectives of this research, setting out what the research aims to achieve: to investigate theory and practice of selection for digital preservation in UK memory institutions. It has defined the scope of the research and described the cultural heritage context in which it is situated. Following this first introductory chapter, the rest of the thesis will be set out as described below.

This research involved a review of the literature, addressing objective 1 in particular. Technical aspects of digital preservation and the concepts involved are explored. In order to identify differences between selection of traditional and digital material it is necessary to establish existing theoretical underpinnings of practice within libraries, archives and museums. This is followed by an examination of particular issues in selection for digital preservation. The findings from this phase of the research are presented in chapter two.

Chapter three explains the research philosophy underpinning the research and the methodology used, including the sampling and analysis techniques employed at each stage.

Chapter four reports the findings from the first set of exploratory interviews with people who have particular expertise in digital preservation, termed 'the experts'. The findings from these interviews form the basis for the second set of interviews with practitioners, i.e. those who currently work in libraries, archives or museums and have some responsibility for selection in their work. This chapter begins to address objectives 2, 3 and 4.

Chapter five of the thesis describes the findings from the interviews with practitioners, addressing objectives 2, 3 and 4. Stakeholders and their roles in selection are identified. Practice is examined in relation to the practitioner as a professional, their knowledge and skills and the role of subjectivity in selection. Their reactions to changes wrought by the requirements of digital material are also described. Relationships with other stakeholders are examined. Selection processes, drivers and influential factors discussed by the practitioners are presented. It also presents the findings from a further examination of current digital preservation policies from a larger sample of institutions which was performed concurrently with the expert interviews.

Chapter six presents the discussion of the findings from the research and the key factors framework, addressing objective 5. The final chapter seven presents the conclusions of the research in relation to the research objectives, along with reflections and recommendations for further work.

## CHAPTER 2 LITERATURE REVIEW

---

The first section of this literature review considers the traditional theory and practice of selection in libraries, archives and museums. Following this the review turns to consider digital material. Technical issues in digital preservation are discussed, providing clarity about the technical problems involved. Definitions of concepts used in the digital preservation literature which are relevant to selection are examined. The final section further examines the literature to focus on factors in selection for preservation of digital material, including criteria, value and policy. This chapter provides an overview and increases understanding of the context and theory of selection for digital preservation within memory institutions in the UK. In this way the chapter will address objective 1.

## 2.1 Selection in memory institutions

This section examines traditional approaches to selection in different domains. The term 'selection' in a memory institution context has synonyms such as appraisal or acquisition and these have specific meanings depending on the context in which they are used. As Ooghe and Moreels (2009) point out, differences in terminology and its uses make cross-institutional comparison and the creation of general selection methods more difficult. These terms will be examined below in the contexts in which they are used and different concepts particular to each domain will be identified.

### 2.1.1 Archival appraisal

In archives selection and appraisal are different processes. Appraisal involves making a judgement on the continuing value of material, so material can be discarded that has no further value (Eastwood 2004, p.202; TNA 2013, p.15) and is most often associated with making value decisions prior to acceptance of responsibility (Craig 2004, p.44). Craig (2004, p.164) describes a two-step system of refinement of decisions, thus defining selection as

*...a second wave or refinement of the initial appraisal decision once the records have crossed the archival threshold....normally selection will take place before records are accessioned into the archive's holdings.*

Similarly Duranti (1994, p.329) distinguishes between 'appraisal for acquisition', i.e. ascribing value before the material is taken in by the archive, and 'appraisal for selection' which happens after acquisition and involves determining which parts of an archive should be preserved long term. Shepherd and Yeo (2003), Craig (2004), and Mercer (2004) give useful accounts of the traditional theory and practice of archival appraisal.

An influential figure in European archival theory and practice was Sir Hilary Jenkinson, an archivist at the UK Public Record Office. Jenkinson believed that the role of the archivist was as a preserver and advisor on records, not as a selector. In order to preserve archives as evidence, the responsibility for appraisal decisions belonged to the creators (Jenkinson 1937, p.149). The decision to destroy records should be taken in the normal course of business by the creators as a record of the transactions of the organisation and the archivist then would take charge of the records that are left when they are no longer actively used. The role of the archivist is to be an advisor and a protector of records (Duranti 1994, p.343) rather than playing an active part; appraisal in this view is seen as 'un-archival' (Cook 2011, p.174). Duranti (1994) echoes Jenkinson's view of the archive as an 'accumulator' of records when she argues that if appraisal is attributing value to records then *'attributing value to that evidence would mean to renounce impartiality, endorse ideology and consciously and arbitrarily alter the societal record.'* (1994, p.344). The Jenkinson view of a passive archivist has been criticised; as Cook (1997, p.24) points out, this approach effectively sanctions the ability of the creator to destroy records

that are potentially damaging to themselves and remove them from public scrutiny. The concept of neutrality is still influential in archival practice, but the archivist clearly makes at least partially subjective decisions on what to keep and what to destroy (Reed 2009, p.124; Cook 2011, p.177). There is a recognition in the literature of the archivist as an active, powerful, participant in the creation of an archive, and so in the creation of societal memory. Hedstrom (2002, p.37) characterises appraisal as an exercise in power:

*...assumptions about impartiality and naturalness obscure the interface of selection and misrepresent the contingent nature of archives. They deny the ways in which appraisal can be both an exercise of power by archivists in shaping social memory and an act of resistance by archivists against other powers that wish to shape social memory for their own purposes.*

The current Code of Conduct devised by the Archives and Records Association (ARA 2012), the professional body for archivists and record managers in the UK, refers only to 'impartiality':

*Members should appraise, select and maintain archival material in its historical, legal and administrative context, thus retaining the principle of provenance, preserving and making evident the original relationships of documents....Members should appraise records impartially basing their judgement on a thorough knowledge of their institution's administrative requirements and acquisitions policies... (ARA 2012, p.2).*

The influence of the Jenkinson view of appraisal can be seen in the UK in the report by the Grigg committee in the 1950s, and in the Public Records Act (1958), where the system by which government records were transferred to the National Archives was devised. It specifies the timing of the transfer of records by instigating a system of two reviews, one at five years and one at twenty five years (Shepherd and Yeo 2003, p.150). The responsibility for appraisal was firmly with a 'reviewer' in the government departments, with TNA providing advice, and this has worked well for paper records (Mercer 2004, p.1). The traditional focus of an archive within an organisation is the keeping of records for evidentiary and legal purposes for a limited amount of time. In the UK the Public Records Act 1958 places a duty on public record bodies to '*make proper arrangements for the identification, safeguarding and transfer of records of historical interest to The National Archives*' (TNA 2012a, p.11). The system has subsequently been revisited due to its inability to manage ever-increasing volumes of records and accommodate digital records, which need appraising much sooner than five or twenty five years (TNA 2012b, p.6).

The method by which the value of records can be assessed is explained by Mercer (2004, p.2), and is based on taxonomy of value devised by the eminent American archivist Schellenberg (1956). Schellenberg (1956) distinguished between records - the larger set of material which is

related to the administration of the organisation- and archives, which is the smaller section which has been chosen for preservation. According to Schellenberg the role of appraisal in modern archives was to determine which records are of potential use to future research. Records were categorised as having primary administrative value to the business and secondary value, the latter divided into evidentiary and informational value (Schellenberg 1956, p.139). Evidentiary value focuses on the importance of the record as evidence, of the creator's actions, functions and policies for example. Informational value concerned the content of records relating to those that are incidental to the original creator, such as other people or businesses (Schellenberg 1956, p.148). Value is also in part dependent on the context of the records; Doom (2004), as part of the draft 'Guidelines on appraisal' for the International Council on Archives (ICA), describes appraisal criteria as depending on context as well as content:

*The production context brings one back to the fundamental archival principle of the 'respect des fonds', a proper respect for the source of archives and for their original arrangement... For the archivist, the production context, bringing together the function and the context within which archives were produced, gives archives their full significance.*

The presence of context and keeping records in their original archival order helps to preserve provenance, i.e. '*the history of ownership related to a group of records or an individual item in a Collection*' (TNA 2011a, p.7) and authenticity, or that the record is what it purports to be (Bearman and Trant 1998), which is important in assessing value. As Eastwood (2004, p.205) makes clear, if there is doubt about the authenticity of records then they are less likely to be selected.

The distinction between records and archives has been an important one, leading to the development of two related professions of records managers and archivists. This distinction is encapsulated in the 'records lifecycle' model (Todd 2006, p.5) where records move through time from creation, to active use, to semi-active use then to disposal, either through destruction or permanent preservation, where they become archives. The view of record keeping and archives as separate functions has been criticised. In the 1990s in Australia a 'records continuum' model was developed, particularly by Upward (1996 and 1997); see Flynn 2001 for an overview of its development. In this model Upward (1996) refers to archives as '*records of continuing value*', de-emphasising the differences between records and archives and consequently between record keepers and archivists. This compares to the lifecycle model, which encourages the division between records keepers and archivists by focusing on a stage based model where archivists are influential in the final stages only where records have changed from being 'active' to being historic. It seems counter-productive to separate records on this basis as historic records are still active, but for a different reason. Here the continuum model is again useful as it sets records in a changing 'space-time', unifying the concept of a record and an archive

regardless of the length of time the material is kept for (Upward 1996; Interpares 2, 2008a p.3). It helpfully regards the keeping of records and archives as a shared responsibility (Upward 1996; Flynn 2001, p.83).

Since the 1950s other approaches to appraisal have been developed in response to the huge increase to records needing to be appraised. The system of macro-appraisal in which the value of items is determined by reference to their function ('functional appraisal') was developed in Canada and is based on the '*... analysis of the creator's key functions, programmes, activities, and interactions with clients, which the records subsequently selected for continued preservation should most succinctly mirror*' (Cook 1997, p.47). It is essentially a 'top down' approach to appraisal which takes account of the context in which records were created. It allows for an analysis of many documents at one time as value is ascribed according to the value of the function they originate from in the organisation (Shepherd and Yeo 2003, p.151). The focus is shifted from the record to the functional context it documents. This approach has many advantages, in particular scalability, as many records can be appraised at once (Bailey 2008, p.111) and it takes into account the context in which they were created. The system helps to determine evidential value resulting from a process and aids consistent decision making across departments. Although functional appraisal is well suited to use within organisations have that well defined or established functions, it is of less use for appraising material for individuals or other social groups that do not have well-defined functions (Craig 2004, p.73). Similar large scale macro approaches have been created, such as the documentation strategy developed in the USA in the 1980s and the 'Minnesota method', also developed in the USA, which take a broad macro appraisal approach. These methods have been criticised because they are labour intensive and have not been widely adopted (Reed 2009, p.125).

Craig (2004, p.49) identifies other factors which may influence appraisal in organisations. These include: the organisation's mission and goals; the role of archives in preserving history or 'the past' and making decisions made now comprehensible to the future; risk management; resources; policies and legal requirements. The TNA (2012b) guidance on disposal scheduling for records managers highlights legal compliance issues when considering disposal of records. These include compliance with the Data Protection Act 1998, which applies to material held by public authorities. There are eight data protection principles (Data Protection Act 1998, Schedule 1) which outline restrictions on the type of material that may be collected and the length of time records may be kept. The Act does include exemptions to the presumption of disposal once the active life of the records is over, which include history, research and statistics (s.33), allowing archives to keep material for these reasons.

### **2.1.2 Selection in libraries**

Selection has long been an issue in libraries. For example Black describes the tension in the beginning of the public library movement between the educational imperative, where only high

quality books should be chosen, and the realistic need to appeal to the public by complying with the demand for recreational fiction (2006, p.35). Gorman and Shep (2006, p. 182-183) suggest that collection building was the aim in libraries until the mid-twentieth century, when it then changed to collection development (Clayton and Gorman 2001, p.xii). Now the term frequently used is collection management. This can be seen as a broad term, including not only collection development activities such as creating policy, user needs assessment and weeding (i.e. deciding what can be deselected or destroyed) but also other processes such as provision of access, evaluation and preservation (Clayton and Gorman 2001, p.xiii) as well as selection. Use of this term recognises that selection is one of a number of related activities (Johnson 2009, p.25; Clayton and Gorman 2001, p.76).

The selection process according to Atkinson (1993 p.98) consists of three stages: discovering or defining the 'anti-collection' (i.e. the set of all publications not held in the local collection); ranking, or deciding on the relative value in comparison to other items, the content of some part of it; and making a decision. Selection decisions are only either 'yes or no' and, as Atkinson acknowledges, the creativity and skill in selection derives from the ranking stage. A more up to date description of selection is found in Johnson (2009, p.109), who describes four stages:

1. Identification of the relevant items through the use of tools such as publishers catalogues or the internet;
2. Evaluation and assessment which asks quality and practical questions about the potential acquisition. This stage includes assessment, where the item is considered against internal criteria.
3. Making a decision to purchase
4. The order is prepared and placed.

Similar to the process described by Atkinson, the decision to purchase and the ordering of material is a largely mechanical operation, assuming the relevant systems are in place. The interest for the librarian lies in the first two stages. It seems the assumption is that once a librarian has decided that an item is of high enough value to the library and its users, the only consideration is 'can we afford it?' Not addressed is that selection can happen at more than one stage in the existence of the material (Feather 2004, p.13), such as weeding, not just at initial acquisition.

The approach to selection in libraries is based on knowledge of a number of contextual factors. Knowledge of users is central to selection of material for a library collection; Bearman (2007, p.27) describes the library approach as selecting items they believe will be of interest to their users, or what Pymm (2006, p.66) terms a 'move from a collection-centred to a client-centred approach'. Clayton and Gorman (2001, p.4), Harvey (2005, p.58) and Johnson (2009, p.108) make clear that selection should be strongly related to user need. Libraries are competing for the attention of users in an increasingly networked environment, which changes user behaviour

and expectations of services, so library processes need to change if they are to continue to meet user needs (Dempsey 2006; Chowdhury 2010, p.217). They are no longer only collecting institutions which provide access, but also focus on facilitating communication and collaboration with and between users (Pymm 2006, p.66; Chowdhury 2010, p.208). The system of patron-driven acquisition (PDA), used in particular in higher education academic libraries, reflects this. This is a 'just in time' collecting strategy based on demand from users; in contrast with the 'just in case' strategy where the library purchases materials in case a user might want it (Breitbach and Lambert 2011, p.17). Books are acquired when a patron expresses a need for them, such as through inter-library loans. A PDA approach can be used to provide access to e-books and other electronic resources, which means that the library does not necessarily have to buy copies of the resource (Shen et al 2011, p.204). However the perspective of librarians and users may be different in terms of the value of the library and its contents, as the information needs of students for example can be satisfied with Google, Flickr, and other Web 2.0 services (Law 2008). The role of librarian as selector may need to change from authoritative to enabling others to select quality material themselves (Law 2008). Furthermore there may be a difficulty in identifying users; large organisations for example may have a very broad heterogeneous set of users and fulfilling the needs of all of these is a difficult task.

In addition to use, knowledge of the mission of the library, of its parent institution and the state of existing collections inform decision making. Clayton and Gorman (2001, p.76-77) suggest that having knowledge of the library purpose is an underlying principle of selection. Some of the questions relating to the existing collection which should be considered include does the item fill a gap in the collection, does the item fall in the scope of the collection development policy or does the item support the institution's interests? (Johnson 2009, p.116). Libraries operating in different contexts, such as academic or public, will have different priorities (Clayton and Gorman 2001, p.80). In academic libraries much of the acquisitions budget is spent on acquiring journals, either in hard copy or access to electronic versions, to support current teaching and research. Business models for journal acquisition allow publishers to influence selection through the library purchasing a subscription to a 'big deal' i.e. access to a specific set of journal titles from a publisher (Taylor-Roe 2009, p.113). The power to select held by libraries is diminished by how publishers choose to bundle their journals (Collier and Van Kiel 2010, p.69). Libraries may purchase journal access in consortia (such as JISC) and this has the advantage of spreading the costs (Woodward and Rowland 2009, p.165). Licensing for bundled 'big deals' are negotiated for nationally by JISC under the NESLi2 licensing scheme (JISC n.d.a), although deals may also be negotiated directly between publishers and libraries. There are advantages for academic libraries in such deals, especially through the NESLi2 scheme, including fixed pricing; a model licence which negates the need for in depth legal expertise; and the use of an intermediary so the library does not have to deal directly with publishers (Collier and Van Kiel 2010, p.71; Cardy and Rumsey 2006, p.70). However these deals may provide access to journals that are not wanted and be limiting; for example Taylor-Roe (2009, p.114)

reports that because her institution spends so much on these deals there is little left to spend on purchasing journals from smaller publishers.

In UK libraries professional ethics and standards have been the subject of much debate over time. The Library Association, which represented the library profession before 2002, began to consider a Code of Ethics in the late 1970s, when faced with potential political problems (Sturges 2003, p.95). Ethical issues of censorship are related to selection; the end result is the same, that some material is not available to users. Censorship and selection are different processes but it is important to make clear the distinction. Asheim (1953) described this difference as where the selector positively tries to find reasons to keep a book in an inclusive way, but the censor negatively tries to find reasons to exclude it. Commentators have found this liberal view attractive, such as Malley (1990, p.28) and Clayton and Gorman (2001, p.77). There is also a professional library ethic to be unbiased or neutral in selection decision making; this is explicit for example in the IFLA (International Federation of Library Associations and Institutions) 'Code of Ethics for Librarians and other Information Workers' (IFLA 2012) and the set of Ethical Principles (CILIP 2012a) and 'Code of Professional Practice for Library and Information Professionals' (CILIP 2012b) developed by CILIP (Chartered Institute of Library and Information Professionals) in the UK. The latter states that librarians should exercise *'Impartiality, and avoidance of inappropriate bias, in acquiring and evaluating information and in mediating it to other information users.'* (CILIP 2012b). The research by Shen et al (2011, p.216) noted in their comparison of selections by users and librarians that:

*Patrons generally search for titles to satisfy a present need, whereas librarians must collect with a longer view of future needs for student learning, faculty research and the preservation of knowledge.*

This demonstrates not only different motivations for selection decisions between users and librarians, but also implies a continued need for librarians with particular knowledge and skills. Clayton and Gorman (2001, p.77) suggest as one of their principles underlying library selection that librarians should *'have the ability to make 'well-informed, independent judgements.... as the competent professional is in the best position to determine the right materials...'*

The current 'Professional knowledge and skills base' framework from CILIP (2012c) includes 'Collection management and development' as one of the key knowledge and skills sets which makes the profession unique (CILIP 2012a, p.3) and this includes selection. Preservation is also included in the framework, but under the 'Records management and archiving' set of knowledge and skills, which also includes digitisation and curation. Whilst the mission of libraries, similarly to museums and archives, may have a preservation element (Maron, Yun and Pickle 2013, p.36), not all library material is collected with an assumption of permanence; that which is to be kept permanently tends to be found in 'special collections' in libraries (Feather 2004, p.13).

Although 'records management and archiving' is seen as part of the overall set of potential skills used in the library, information and knowledge professions, looking after material in the long term in libraries is more analogous to the activities and role of archivists (Feather 2004, p.12). Although libraries have been closely involved in digital preservation from the outset, the literature demonstrates that archival theory underpins much research and practice. Digital libraries, defined as '*An organisation, which might be virtual, that comprehensively collects, manages and preserves for the long term rich digital content*' (Candela et al 2007, p.17), include preservation as an activity, but the underlying archival theory is made explicit by Ross (2012, p.50):

*...digital libraries are more akin to archives than they are to conventional libraries, we need to seek their theoretical foundations in the domain of archival science and their practices in archival and records management environments. Archival science, with its principles of uniqueness, provenance, arrangement, and description, authenticity, appraisal, and its tool sets such as diplomatics and palaeography, may offer us a framework for a theoretical foundation for digital libraries.*

### **2.1.3 Selection in museums**

The museum literature focuses on the process of acquisition. This is defined by the Museums Association (MA) as '*the process of obtaining legal title to an item with the intention of using it for museum purposes*' (MA 2004, p.2). These purposes are found in the definition of a museum; ICOM (International Council of Museums) (2012) define this as an institution that:

*...acquires, conserves, researches, communicates and exhibits the tangible and intangible heritage of humanity and its environment for the purposes of education, study and enjoyment.*

Modern museums developed in Europe in the eighteenth century from the collections of private individuals; the Ashmolean museum was founded in 1683 based on a collection given by Elias Ashmole to the University of Oxford for example (Ambrose and Paine 1993, p.6). Museums still acquire much of their collections from donations or gifts, but also bequests, purchases, exchanges or field collection, (MA 2004, section 6; Ladkin 2004, p.20). As the methods of acquiring items for museum collections have been ad hoc and non-standard, so has collection building.

Traditionally the museum is seen as a 'sacred space' where the museum is a 'shrine' with objects in it that offer enlightenment and a transformative experience (Marstine 2005, p.9). The museum protects the objects, they are 'treasures', looked after by curators. In this view visitors are there to be given information, to be educated, and are not seen as co-creators or having a role in interpretation. This reflects the view found in the wider heritage literature of the assessment of significance as an act of power (Smith 2006, p.29; Gibson 2009, p.75-76).

Marstine (2005, p.5) also gives other views of a museum, including 'new museum theory', which recognises that although museum workers 'naturalise' their policies and procedures as professional practice, the decisions that are made by curators are positioned in particular value systems. The museum literature recognises that 'neutrality' or a lack of bias in decision making, similar to that in libraries and archives, has not been relevant to curators; as Young (2006, p.189) points out, traditionally individual curators use their own judgement when selecting what to collect:

*More often, acquisitions are the product of the judgements of connoisseurship (whether or not consciously claimed as such). This is the most traditional strategy of assessment, whereby the curator applies his or her knowledge to determine which objects are relevant and meaningful to the purposes of the collection*

The 'Code of Ethics for Museums' from the MA only discusses acquisition in terms of '*Acquire items honestly and responsibly*' which includes '*Acquire an item only after thorough consideration of its long-term value and how it will be used*' (MA 2008, p.14). The interest of a museum or a department in museum in engaging with digital material is in part dependent on the influence and enthusiasm from senior curators or managers (Hudson 2012, p.46), which seems to reflect the individual exercise of judgement mentioned by Young.

Once an item has been acquired it is 'accessioned', defined by the Collection's Trust (Harrison 2005, p.1) as:

*The formal addition of an object to an institution's collections. It follows transfer of title and includes assigning an accession number to an object and recording details in an accessions register.*

Some items will be acquired for use by the museum, such as for handling or demonstration purposes, without being formally accessioned. It is a formal process that should not be undertaken without much consideration as it demonstrates that responsibility for the object is taken permanently and it will be cared for in perpetuity (MA 2004, p.2).

There is an ethical dimension to museum collecting, exemplified in the MA's 'Code of Ethics' (2008) which contains clauses that have a bearing on selection such as '*Acquire items honestly and responsibly*' (MA 2008, p.14), referring to collecting according to published policies (clause 5.1), consideration of long term value (clause 5.2) and exercising due diligence on provenance and ownership (clause 5.7). It also states that '*museums hold collections in trust on behalf of society*' and that '*Museums behave as ethical guardians as well as owners of collections*' (2008, p.10), underlining the permanent nature of museum collections. In a deliberately provocative piece, Knell (2004, p.19) argues that the focus in museums is too much on collecting for

collecting's sake, and museums should question their assumption of permanent responsibility. He convincingly argues that some items in museums are collected for nostalgia or because it is assumed that no-one else will. These then take resources to preserve but the underlying purpose of collecting such objects is unclear or untenable. Museums should question their collecting activities by asking why it is important that particular objects are collected and how long they should be kept for (Kneel 2004, p.19), reflecting the view of the MA to collect in a responsible manner.

Standards and guides for collecting and managing acquisition have been developed, despite the practicality of these being questioned; Pymm (2006, p.63-64) for example asks if it is possible in a museum context which has to accommodate significance that is geographically and temporally based. Despite such arguments, guides to collecting include the PAS197:2009 (Winsor 2009) and the 'SPECTRUM 4.0' standard (Collections Trust 2011), both of which were created by the Collections Trust and include collection management processes and workflows. These recommend that collecting be controlled by collecting policies, as does the ICOM (International Council of Museums) Code of Ethics for Museums (ICOM 2006, p.3). The Museums Association (MA 2004, p.3; 2008, p.14) gives criteria that should be considered when selecting an item. The Collections Council of Australia's publication 'Significance 2.0: a guide to assessing the significance of collections' (Russell and Winkworth 2009) attempts to provide guidance to librarians, and others working with collections, on value and significance. It suggests four primary criteria that may be used to assess the significance of collections, including historic, artistic or aesthetic, potential research value and social or spiritual. Four further criteria are suggested, which modify the first four: provenance; rarity or representativeness; condition or completeness; and interpretative capacity (Russell and Winkworth 2009, p.20). The guide makes the benefits of assessing significance clear, including aiding selection for preservation actions and collaboration between institutions. The Significance 2.0 framework is also interesting as it attempts to give overall guidance on assessing significance for any type of collections held in archives, libraries, museums and galleries, whilst clearly acknowledging the different approaches of each domain (2009, p.3). However this approach does not seem to have had the same impact on the archive or library domains, despite the aim of the framework to apply to all. However their approach has been influential in the UK in the museums domain. The authors of the Renaissance East Midlands (REM) 'Reviewing Significance: a framework for assessing museum collections' significance, management and use' (2010) guide explicitly acknowledge the influence of the Australian framework (2010, p.4). The REM guidance aims to help museums assess the significance of items and whole or part collections which they already own.

#### **2.1.4 Summary**

In this section, traditional approaches to selection and appraisal in libraries, archives and museums have been discussed and several themes have become apparent. Firstly it is clear

that each domain has different roles which affect their approaches to selection. There is recognition within the museums sector that collecting in the past has been ad-hoc; this may reflect their origins as collections from individuals and their particular use of donations as a method of acquisition (Ambrose and Paine 1993, p.125). Systematic guidelines for assessing significance of items have been developed; these help to counteract the criticisms found in the wider heritage literature that collecting may have been biased in the past. Archival appraisal is based strongly within a traditional document and organisational context and has a body of underpinning theory on which practice is based. Libraries on the other hand focus more on collecting for current use and providing access to material for their users. Criteria identified include costs, the mission of the library and the composition of the current collection. Whilst there is little in the library literature that addresses selection theory, a number of underlying principles can be traced (Clayton and Gorman 2001, p76) which include knowledge of the current collection and knowledge of potential sources of material to select from.

Selection is clearly a process that is performed at different stages. In museums there is a two stage process of acquisitions and accessioning where permanent responsibility is taken for items when they are accepted into the permanent collection. In archives it is also recognised that selection and appraisal can happen both before and after acquisition (Craig 2004, p.164), though traditional archive theory states that it should have been performed by creators prior to acceptance by the archive. In libraries selection is viewed as a process in the overall function of collection management.

The literature also reveals that assumptions are made about taking responsibility for items once they are collected. There often is the assumption that this responsibility will be permanent, once the items are in the institutions collections; ownership of a physical item implies permanent responsibility, especially in museums. Exceptions to this are in records management where there is an assumption of disposal, that items will be kept only for a certain length of time and that the majority of documents will be disposed of before they reach the archive, once they have come to the end of their useful life. Once the items are in the archive then there is a stronger assumption of permanence. There is also an exception in the museum context where items that have not (yet) been accessioned or have been collected to be used in a particular way such as for handling or demonstrations are not part of the permanent collections. Museums and libraries may collect for different practical purposes e.g. handling, reference, or lending. Libraries also have non-permanent collections, such as where the users require up to date information and so collections are weeded to remove out of date publications. This contrasts with 'special' collections in which are the 'treasures' of a library, which may be archival in nature and are managed and preserved separately. The general assumption of responsibility places institutions in a difficult position where collections continue to grow but little is disposed of; questions of resources and space then become more important.

Archival theory based on Jenkinson (1922) states that archivists should be neutral collectors of material, though it is clear that this is not the case in practice; there is a recognition that archivists are active participants in the creation of an archive (Hedstrom 2002, p.37). The records continuum theory (Upward 1996) positions archivists and record keepers in the same space, in contrast to the lifecycle theory in which archivists are only influential once the records have ended their active life. Neutrality is an ideal in libraries too, though it is framed ethically as demonstrating a lack of bias when collecting. The issue of neutrality has not been as prominent in museum literature, where in the past the curator has been viewed as a 'connoisseur' who uses their own knowledge to select items.

Archives have already seen an increase in the volume of records and have adapted theory and practices to manage this material, such as functional or macro methods of appraisal (Cook 1997, p.47). Appraisal is based on a taxonomy of value relating to the records role as evidence; greater value comes from the presence of contextual information for the records. The library literature focuses more on practical guidance rather than developing theory, though librarians are seen as having particular collection management skills. Museum emphasis on the care of individual items and has not focused on dealing with a volume of material, although it is recognised that they hold large collections of material not on display or in use. The museum literature gives practical guidance on assessing significance after the material is acquired, although some practical criteria have been suggested (e.g. Russell and Winkworth 2009). The next section focuses on the technical aspects of digital material and relevant concepts found in the digital preservation literature.

## 2.2 Digital preservation

Traditional material can be preserved 'passively' by providing the appropriate storage environment and controlling environmental variables such as temperature, humidity and pests (IFLA 1998). For digital material it may be possible to preserve material passively by storing the bit stream in an appropriate place, perhaps until workflows or policies have been created to manage it (Paradigm 2008), but it is more likely that digital material will require an active program of preservation to ensure it is continually accessible and useable. This difference is described by the DPC Handbook (2008, p.25) as:

*...the need actively to manage the resource at each stage of its life-cycle and to recognise the interdependencies between each stage and commence preservation activities as early as practicable. This represents a major difference with most traditional preservation, where management is largely passive until detailed conservation work is required, typically, many years after creation and rarely, if ever, involving the creator.*

The technical aspects of digital material that prompt this need for early interventions will be examined in the next section.

### 2.2.1 Technical aspects

Preserving digital material is more complex than preserving traditional forms of heritage, so it is unsurprising that much research has addressed technical issues in preservation. Digital data is susceptible to loss, and two issues in particular have been focused on: media instability and technological obsolescence. The media on which data is stored is fragile and unstable; even assuming optimal storage conditions, magnetic devices and optical storage such as CDs and DVDs have a short expected lifespan. A second problem arises from the technological obsolescence of the hardware which is used to access the data, or the software needed to read the data. Changes to formats, coding, operating systems and other aspects can render information unusable. The threat of technological obsolescence is a greater threat than deterioration of physical media, which may outlast the hardware and software needed to use it. This problem was recognised early in the digital preservation literature, see, for example, Fresko (1995) and Waters and Garrett (1996, p.5). Rothenberg (1999, p.2) summarises the technical problems:

*It is now generally recognized that the physical lifetimes of digital storage media are often surprisingly short, requiring information to be “refreshed” by copying it onto new media with disturbing frequency. The technological obsolescence of these media (and of the hardware and software necessary to read them) poses a different and equally urgent threat. Moreover, most digital documents and artefacts exist only in encoded form, requiring specific software to bring their bit streams to life and make them truly usable; as these programs (or the hardware/software environments in which they run) become obsolete, the digital documents that depend on them become unreadable—held hostage to their own encoding.*

There have been a number of strategies suggested for preserving digital material. One option for preserving digital material is to keep the technology working, so old computers for example would be kept to run obsolete software on. By using this strategy the look, feel and experience of the digital material would be kept. Problems with this approach include the physical deterioration of hardware and the disappearance of expertise to operate the systems. Thibodeau (2002, p.17) and Stawowczyk Long (2009, p.7) describe this approach as unsustainable in anything but the short term and too complex though Anderson, Delve and Pinchbeck (2010, p.111-112) argue that this view is not necessarily shared by all communities outside of digital preservation and that ‘computer museums’ have a continuing role (2010, p.113). It could be still appropriate in certain situations such as museums which preserve computer hardware, or in preserving computer games (Barwick, Dearnley and Muir 2008, p.6).

Two further strategies have received more attention in the literature. Migration involves copying the data from one hardware or software generation to a newer one as they become obsolete

(for an overview see Pearson and del Pozo 2009, p.9). Despite its relative popularity as a strategy (DPC Handbook 2008, p.112) and potential for automation (Ferreira et al 2006), there are associated problems. Strodl et al (2007, p.31) for example mentions that each migration performed contains risks to the data and as the strategy changes the data it can only preserve some of the characteristics of the material. So particular fonts may not be available in a new word processing format for instance. Anderson, Delve and Pinchbeck (2010, p.114) also point out that there is a question of scale and practicability; it is logically possible to migrate material to new platforms when each item requires it, but this is not practical due to the large amount or complexity of material that may need migrating. Another approach to migration is to migrate objects to a standard format when they are first accessioned. This approach is known as 'normalisation'. The advantages of the normalisation approach is that it postpones the need to migrate information and when migration is necessary there are far fewer formats to migrate (Rosenthal et al, 2005). However this does not solve the problems inherent in migration in the long term as it only postpones them and there is no guarantee that the formats used will not also become obsolete.

An alternative strategy is to emulate, or recreate, the hardware or software environment of the digital object. Strodl et al (2007, p.30) define it as:

*the duplication of the functionality of systems, be it software, hardware parts, or legacy computer systems as a whole, needed to display, access, or edit a certain document., aiming to retain the functionality of the digital object.*

The EU-funded KEEP (Keeping Emulation Environments Portable) project newsletter (Mikolaskova2010, p.1) explains advantages of emulation as leaving the original object untouched; no periodic migration is needed and the appearance and functionality of the original are kept. Stawowczyk Long (2009) describes work in the National Library of Australia in testing a variety of emulation and migration tools to preserve web archives. All emulators tested were found to be very complex and time consuming to configure (2009, p.44) and Stawowczyk Long concluded that none of the tools tested provide a solution to the question of preservation, so further work is necessary (2009, p.47). The KEEP project has taken forward research into emulators, taking an approach that allows migration of an emulator, which is a digital object, rather than creating them from scratch each time there is a need by focusing on creating a 'virtual machine' which can itself be migrated (Anderson, Delve and Pinchbeck 2010, p.115).

Whilst much digital preservation literature reflects Rothenberg (1999) in focusing on the risk to data from software format obsolescence (Waller and Sharpe 2006, p.8; DPC Handbook 2008, p.36; del Pozo, Stawowczyk Long and Pearson 2010, p.292 for example), this focus has been questioned. Rusbridge (2006) and Rosenthal (2010) ask whether formats are changing as quickly and irrevocably as is assumed. Rosenthal argues that format obsolescence is a rare

occurrence as the maturing market for software leads to the dominant software being supported by the 'winning' company to encourage others to use their products and to keep their dominant position, such as with Microsoft Office products (2010, p.2). Backwards compatibility needs to be built into new versions of such products otherwise people would not pay to upgrade. He also argues that the change from paper to Web publishing for e-documents has reduced the incentive for making formats obsolete as the author wants as many people as possible to read their document, and the costs of maintaining formats via a web browser is minimal (Rosenthal 2010, p.3). There are other, perhaps more pressing, threats to digital material than format obsolescence; Rosenthal et al (2005) and Barateiro et al (2010, p.9) identify a broad taxonomy of threats to digital material, including not only obsolescence but also faults with hardware, software and media; infrastructure or communication faults and failures; disasters, both natural and from human error; attacks; economic and organisational failure; and legal changes and requirements.

The focus in digital preservation research has partly moved away from obsolescence toward other issues. For example, digital preservation research funded by, or of potential interest to, the European Union (EU) more recently centres on issues such as automation of preservation functions, the preservation of complex objects and processes, building networks of digital preservation experts, training for digital preservation and the development of auditing and certification schemes (Strodl, Petrov and Rauber 2011; Billenness 2011). Attention has moved to preserving the context of digital objects, based on concepts derived from the OAIS reference model. This model was developed initially for space data by the Consultative Committee for Space Data Systems (CCSDS). It has been used widely to underpin digital preservation research, the development of digital archives and other digital preservation standards (Ruusalepp et al 2012, p.120). Digital preservation is reliant on interoperability between systems and so is dependent on standards (Ruusalepp et al 2012, p.117). There are many British and International standards which relate to different aspects of the preservation of digital material and lists can be found in the Interpares 3 project (Xie 2007), the DCC DIFFUSE (Dissemination of InFormal and Formal Useful Specifications and Experiences) project 'standards frameworks' (2009), TNA's 'Framework of Standards' (2012c) and Ruusalepp et al (2012). These are not exhaustive however; Ruusalepp et al (2012, p.117) estimate that there are at least two hundred standards relating to digital curation and preservation. These are relevant to many aspects, including metadata, such as the Preservation Metadata Implementation Strategies (PREMIS) data dictionary (2011); file formats such as how to use Portable Document Format (PDF) for preservation (ISO 19005-1:2005); and ISO 14721:2003, the Open Archival Information System (OAIS) reference model.

The OAIS reference model describes an archive, which is an organisation that has accepted responsibility for preserving information and making it available (CCSDS 2012, p.1-1). There is a separate standard describing 'pre-ingest' activities, which details the relationship between the

creator and the archive prior to ingest of the material (CCSDS 2004). In the OAIS model the digital (or data) object is 'An object composed of a set of bit sequences' which then needs additional 'representation information' (RI), or 'information that maps a Data Object into more meaningful concepts', such as associated software, to transform it into an 'information object' which can be understood and used (CCSDS 2012, p.2-3). CCSDS (2012, p.1-15) give as an example illustrating this:

*rendering the JPEG file as bits is not very meaningful to humans but the software, which embodies an understanding of the JPEG standard, maps the bits into pixels which can then be rendered as an image for human viewing.*

The RI necessary to understand a digital object may be part of a network of necessary information as the RI itself is an object; in order to understand RI we need further information. To limit the amount of information needed and the RI network, the OAIS model proposes the concept of a 'designated community'. Entities that the system interacts with are characterised in different ways, including as consumers who are expected to use the material preserved by the system, producers of material and those that manage the material (CCSDS 2012, p.2.2). A subset of consumers is termed the 'designated community' which are those consumers expected to be able to independently understand the information (CCSDS 2012, p.1.11). The designated community has an associated 'knowledge base', which is the knowledge it uses to understand RI (Patel et al 2009, p.213; Chowdhury 2010, p.215). This concept enables a group of stakeholders to be identified who can be used to test the success of preservation activities if they still understand it (Giaretta 2011, p.16).

### **2.2.2 Significant properties**

The choice between preservation strategies relies in part on the properties of the object that the institution wants to preserve. The 'CURL exemplars in Digital Archives' (CEDARS) project described the concept of 'significant properties' to depict these. It describes significant properties as the content and functionality that is required by the archive (2001, p. 14). Since then the definition has matured; the 'Investigating Significant Properties of Electronic Content' (InSPECT) project (Grace, Knight and Montague 2009, p.3), explains significant properties as:

*The characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record.*

The purpose of identifying significant properties is also made clear by Wilson (2007, p.7):

*Unless such properties can be defined in a rigorous and measurable manner, cultural memory institutions have no objective framework for identifying, implementing, and validating appropriate preservation strategies, nor for asserting the continued authenticity of their digital collections.*

Essentially the purpose of defining significant properties is so that it is clear whether a preservation action has been successful and that the material is still authentic. Authenticity is an archival concept meaning that the record is what it purports to be (Bearman and Trant 1998; Wilson 2007, p.4); the DPC Handbook (2008, p. 24) defines it as '*whatever is being cited is the same as it was when it was first created unless the accompanying metadata indicates any changes.*' Having confidence in the authenticity of records, whether paper or electronic, is seen as important, but especially so for digital where the material may be changed easily and is so when material is migrated from one environment to another. The criteria for authenticity will vary between stakeholders, according to their context, so for example there is a strict authenticity requirement for those using legal documents than for other stakeholders, where the requirement may be for only the intellectual content (Knight and Pennock 2008, p. 162). The underlying assumption is that it is possible to determine an authentic final version of something, which for non-digital material need not be problematic; for digital material that is dynamic and easily changed it can be. Hence the determination that digital objects do not need to remain in an unchanged state in order for them to be considered authentic (Wilson 2007, p.4); what is necessary is for the changes to be documented in the accompanying metadata and for all the transformations that the material has been subject to throughout its lifecycle are traceable (DPC 2008, p.24; Salza et al 2012, p.26). This means that selection of digital material should take account of the high requirement for appropriate metadata and contextual information to be selected or created for the material.

The determination of significant properties is widely considered important in the digital preservation literature (Knight and Pennock 2008, p.160; Wilson 2007, p.7). The term significant properties is well-used in the literature, but is recognised as problematic; '*The term significant characteristic has become over-loaded and remains ill-defined*' (Dappert and Farquhar 2009, p.298). Other terms for similar concepts may be used; the National Archives of Australia have used 'essences' to describe significant properties (Heslop, Davis and Wilson 2002, p.13) and 'essential characteristics' is another similar term (Grace, Knight and Montague 2009, p.4). Additionally there are different meanings of 'significant' in 'significant properties'; significance and value, which are closely related, are often used interchangeably to indicate a measure of perceived worth, which may cause confusion. Also the level at which significance is to be assessed is unclear - it may refer to 'information-level' or technical properties of a file that need to be preserved or in a broader sense which includes content (del Pozo, Stawowczyk Long and Pearson 2010, p.293). Moreover, the concept is based on the assumption that a digital object has immutable and discernible information content that is to be preserved and this is not always

so (Anderson, Delve and Pinchbeck 2010, p.116). Value in particular changes over time. Value plays a role in selecting significant properties, and Knight and Pennock (2008, p.161) turn to an archival definition which relates to intrinsic value - the value of something for itself – and ‘extrinsic’ value, that is value related to an external function. Determination of value relies on an archivist or collection manager selecting through making subjective value judgements based on sets of criteria (Grace, Knight and Montague 2009, p.4). Each stakeholder will have different criteria for evaluating authenticity, which is influenced by the context of their work. Different institutions will have different requirements for which significant properties are most important for them. The collection manager must make decisions of what it is that preservation action is attempting to preserve. Is it the look and feel of the object? Or is it the functionality? (CEDARS 2002, p. 15).

Difficulties with the term ‘significant properties’ are avoided by del Pozo, Stawowczyk Long and Pearson (2010, p.295-296) when they discuss the concept of ‘preservation intent’. They explore the idea of a rendition of a file as a performance i.e. a version of the object that the user sees. The implication is that it is the performance that may be important, rather than the object itself, so it would not matter if the source or process used to render it is changed. They point out that to most archivists changing the source would not be acceptable, even if the process is changed, as they value authenticity. They also argue that what is considered an adequate representation is subjective; representations may not be possible that others would find more appropriate if the object is changed. One strategy would be to define the significant properties of the item, but as del Pozo, Stawowczyk Long and Pearson (2010, p.293) point out, this tends to focus on the information level properties of a file; this they describe as:

*The binary form of a digital object is a means of encoding information, such as specific strings of characters or numerical values, which might represent anything from the name of a photographer to the colour value of a single pixel.*

They suggest a similar though broader and more nuanced strategy of focusing on ‘preservation intents’ (2010, p.297), which includes assessing the need to preserve the item, the specific characteristics of the item, which may include properties of the hardware, bitstream, information and performance, and also the time frame that the institution needs to preserve the item for. ‘Preservation intent’ therefore captures different preservation goals and abilities to which may vary between items and institutions. Similarly to defining significant properties, del Pozo, Stawowczyk Long and Pearson (2010, p.299) argue that by articulating the terms of preservation, auditing preservation actions is easier. This approach seems more inclusive, focusing on more aspects of digital material than the concept of significant properties as used in the literature.

### 2.2.3 Summary

This section has focused on examining the technical issues relevant to digital material and digital preservation as well as relevant conceptual issues found in the digital preservation literature. Since the 1990s research has been ongoing into how to preserve digital material. It is clear that it has different properties to traditional material that mean it needs more active preservation management earlier in its lifecycle. The media on which it is stored and the hardware and software used to render it have been a focus of research, with the underlying assumption that these are under threat from obsolescence. Strategies such as migration and emulation have been devised to preserve digital material and much research has been conducted into these technical aspects. More recently the threat from technological obsolescence has been viewed as perhaps less immediate than other issues, such as how to preserve context along with the object; research is ongoing and further work is necessary to preserve complex digital objects. That preservation is an ongoing process is clear but the terminology used is imprecise. Material is not necessarily 'preserved' as an object but it is the ability to continue to use the material that is important. Focusing on authenticity in terms of an unchanging object is therefore unhelpful. The concept of significant properties is used in the digital preservation literature to refer to those aspects of the material which need to be preserved over time. This concept is linked to selection though the use of the term significant; using this term to refer to specific properties and its use to mean 'value' is confusing and leads to a lack of clarity.

The next section will examine the literature focusing on selection and includes issues relating to the management of digital preservation.

## 2.3 Selection of digital material for preservation

Selection for digital preservation is a necessary part of the process of preservation (DPC Handbook 2008, p.103). As is made plain by Lunghi et al (2012, p.201):

*...there is alignment about the value - indeed the necessity - of selecting and appraising digital information: in effect, assigning value to it and prioritizing some data as more valuable than others. There is, however, less alignment about the practicality and processes for actually carrying out selection and appraisal routines.*

Whilst the necessity of selecting material for preservation is common throughout the digital preservation literature, agreement on the processes and procedures to be followed is not, reflecting the different traditions and approaches found in different memory institutions. Digital material has particular requirements and properties that mean that traditional forms of selection or appraisal may not be appropriate (Russell, 1999; Eastwood 2004, p.202; Bailey 2008, p.72 for example). These include increased volume of material, multiple copies or versions of the

same item, and being stored on older media so the material needs processing prior to appraisal (Paradigm 2008). Further factors which may affect selection of digital material for preservation as discussed in the literature will be explored in the rest of this chapter.

Different strategies for selection are apparent in the literature when choosing material to preserve from the World Wide Web is considered. JISC and the Wellcome Trust commissioned research in 2003 on this topic (Day 2003). Whilst it is now ten years old this research provides a useful introduction to the issues of web archiving and some of the relevant strategies and initiatives. Day (2003, p.7) describes problems of the Web's quickly changing and 'fluid' nature; unclear responsibilities for preservation; and copyright issues. The JISC-funded PoWR (Preservation of Web Resources) project examined the preservation of web resources, including selection and appraisal strategies. The resulting 'Guide to Web Preservation' (Farrell 2010, p.21) describes three such strategies: bulk or domain harvesting; criteria-based selection; and an event based selective approach. The approaches taken by national libraries to selecting web material for preservation demonstrate both selective and non-selective strategies to manage selection on a large scale. Periodic harvesting of the whole web domain is performed for example by the National Library of Sweden using automated harvesting robots. The selective approach is used by national institutions such as Australia's Preserving and Accessing Networked Documentary Resources of Australia (PANDORA) archive and the National Library of Denmark. The selective approach involves using a predetermined set of criteria to choose which websites to preserve, perhaps based on specific themes or events. In the UK the British Library (BL) and partners including the National Library of Wales, the Wellcome Library and JISC provide the UK Web Archive. Here staff choose specific web material for preservation, according to specific criteria, allowing contributing institutions and individuals to nominate websites for inclusion whilst retaining final curatorial control. The selective approach may have a number of advantages, including limiting the amount of web material collected, increasing quality control, aiding cataloguing and making the material easier to manage (ULCC and UKOLN 2008, p.19; Phillips 2009), although the potential for selector bias has been acknowledged (Pennock 2013, p.10). The automated and selective approaches may also be used in conjunction, such as by the Bibliothèque nationale de France (BnF), who performs both bulk automatic harvesting of French websites and focused crawls based around themes or events of sites chosen by library staff (BnF 2011).

### **2.3.1 Criteria**

Criteria provide a method of articulating value, and allow comparative assessments to be made between competing material. Criteria for selection have long been an issue in the literature regardless of format. A taxonomy of preservation for microfilming was suggested by Atkinson (1986) who divided material into different classes based on the type of value they have, such as economic or high-use. Despite the debateable nature of the classes, Atkinson recognised different motivations for preservation, such as protecting the capital value of the collection.

Further research which includes selection criteria has been conducted in the context of selecting analogue material, including books and microfilms, to be digitised (Ayris 1998; Hazen et al 1998; Edwards, Matthews and Nankivell 2000; NINCH Guide 2002 for example). Gould and Ebdon (1999, p.12) surveyed national libraries, universities, archives and other cultural heritage institutions and found that the most prevalent criteria were (in decreasing order):

- historical and cultural value
- to increase access
- academic importance
- to reduce damage
- preservation.

Over half of all respondents chose these criteria, as opposed to 15% or fewer who chose 'save space', 'research in to digital processes' or 'commercial exploitation'. Although the question included an option to describe any other criteria used, it is not clear from the report if any more were suggested. The work by the DISCmap project (Birrell et al 2009) also focuses on criteria for digitisation, specifically of special collections in a higher education context, but interestingly compares user priorities with those of '*intermediaries (librarians, archivists and curators)*'. They highlight similarities and differences in criteria considered appropriate by each group; this highlights a difficulty with criteria whereby assumptions made by those creating and using criteria about what their audience values may not be accurate. A study by Ooghe and Moreels (2009) analysed criteria for digitisation obtained from policy documents in institutions from various countries. They found many criteria which they grouped into six categories: institutional frameworks; value; physical criteria; copies and multiples; metadata and financial frameworks. Whilst some of the criteria are similar to that proposed in earlier work, Ooghe and Moreels (2009) reveal a greater emphasis on issues relating to long term sustainability, though specifically focussed on criteria for digitisation, deliberately setting to one side the problem of born digital material and long term preservation. This is in contrast to the findings from the PORTICO report (2011), which considered the activities of institutions in preserving the digital output from digitisation projects, finding that:

*Cultural heritage organizations do not often have a sustainability plan associated with their digital content. Rather than considering the digital content to be a product that must be sustained, it is considered another outlet for their special collections (2011, p.18).*

As Edwards, Matthews and Nankivell (2000, p.21) note, selection for digitisation has different motivations and implications than selection for preservation and it seems that often digitised versions of physical material are treated and valued differently than born digital material.

Despite a small sample size, an interesting survey of preservation activity in archives in Wales found that 36% of respondents did not seem to have any criteria and they note that those that did seemed to base criteria on analogue material (McInnes and Phillips 2009, p.5). A recommendation from this report is to highlight the need to update policies with selection criteria suitable for digital preservation though the report does not say what these should be. It seems that few criteria have been formulated specifically for preservation of digital material; the assumption is that the same criteria will be appropriate as that used for traditional material. There is little in the literature examining criteria specifically for preservation, but what there is highlights the importance of an institution having the necessary ability to preserve digital material, both technologically and economically (Seadle 2004). The Decision Tree developed by the DPC (2006) was created to help organisations formulate selection policy and includes criteria which institutions might consider in selecting digital material for preservation. Those in the first section, which focuses on 'selection of version and content', include institutional mission and long term value; these are similar to criteria used in selecting analogue material. The Decision Tree then goes on to further sections which consider rights and responsibilities; technical and cost issues; and documentation and metadata. The Decision Tree focuses more on strategic factors, reflecting its intended use as underpinning policy but does not directly address some drivers which are important, such as access and demand. The advice given within the text however does mention further considerations, such as '*Do you have or can you negotiate acceptable access arrangements?*' Also some of the questions may be difficult for an institution to answer, such as '*is it in a format you can manage now or in the future?*' The guidance provided by Whyte and Wilson (2010) includes suggested criteria, and these are:

- relevance to mission
- scientific or historic value
- uniqueness
- potential for redistribution (i.e. IPR issues, acceptable formats, appropriate storage to ensure its integrity, control of sensitive information)
- non-replicability (of the data source)
- justifiable costs
- full documentation.

These criteria are specific to the research data context which the guidance is aimed at and so some of these criteria are not necessarily directly relevant to other contexts, such as 'non-replicability' (which seems similar to uniqueness). There are issues with their list; for example 'full documentation' includes metadata, but this is also considered in the 'potential for redistribution' criterion, which has many sub-categories that should be separate. Harvey (2007, p.10) also discusses criteria for preservation; for digital material the emphasis is placed more on technical capability, costs, and the need to make earlier decisions along with a greater

importance of deciding what aspects of the material to preserve and legal complexities than with traditional material. He clearly states (2007, p.12) that:

*A universally-applicable appraisal and selection framework for data is not realistic. Different kinds of digital materials, created in different contexts for different stakeholders, require different approaches to appraisal and selection.*

However he then attempts to do this (2007, p.37). Through examination of the criteria suggested by Harvey (2007), Whyte and Wilson (2010) and in the Decision Tree (2006), along with criteria mentioned in the preceding literature, it is possible to formulate a 'core' set of criteria that are most commonly used to select for either digitisation, or preservation. Table 1 lists the core criteria, along with where each has been highlighted (those sources highlighted in green focus on digitisation):

Table 1 Criteria

	Ebdon and Gould (1999)	Ooghe and Moreels (2009)	DISCmap project (Birrell et al 2009)	Harvey (2007)	DPC Decision Tree (2006)	Whyte and Wilson (2010)
Value	X	X	X		X	X
Costs		X		X	X	X
Collection policy or mission		X		X	X	X
Documentation		X		X	X	X
Access	X	X	X			
Demand		X	X			
Intellectual property rights (IPR)		X		X	X	X
Condition of original	X	X				
Commercial exploitation	X	X				
Use	X	X	X	X	X	
Technical ability/infrastructure				X	X	X

The table reveals that there is agreement in the literature that core criteria for digital preservation are: costs, collecting policy or mission, the existence of documentation, the issue of intellectual property rights, and technical ability and infrastructure. Whilst there are parallels between criteria suggested for digitisation, such as for use and value, the criteria suggested from the literature focusing more on preservation include technical considerations which

assume much greater importance, including format and the ability to manage different types of material. Also recognised as important are the ownership or negotiation of intellectual property rights. Whilst this is pragmatic it does mean that items of value may not be selected for preservation because they are not easy to preserve. Table 1 shows that a common criterion for digitisation and preservation is use. Selecting for use is a fundamental criteria for selection (Clayton and Gorman 2001; Harvey 2005, p.58; and Johnson 2009, p.108). The emphasis is on selecting for use now; high recurring value may be equated with access and high use (Russell 1999; Ross 2007, p.3). However if selection criteria for digital material are formulated too much on the basis of current policies and priorities, this could lead to the loss of valuable data (Russell 1999). There is a need to include material not only of current high use but also have potential future use as a criterion (Lunghi et al 2012, p.218). The table also highlights the importance of considering costs and the need to align collecting with the institutional mission and policies; these factors are related as it is more likely that financial outlay will be met if the proposed work clearly meets or enhances institutional aims (Currall, McKinney and Keenan 2007, p.8). A further criterion common in the literature and highlighted in table 1 is value, and as this is a complex concept, it will be examined more closely in the next section.

### 2.3.2 Value

The concept of value is central to the question of selection, as selection is the process of determining what is valuable enough to provide resources to keep. Article 7 of the UNESCO Charter on the Preservation of the Digital Heritage (2003) states:

*As with all documentary heritage, selection principles may vary between countries, although the main criteria for deciding what digital materials to keep would be their significance and lasting cultural, scientific, evidential or other value.*

This does not provide much guidance in determining what 'significance' or 'value' is and there are many possible perspectives. The related concept of significance has been discussed in a digital heritage context by Pymm (2006) and Lloyd (2007, p.54) who argue that assigning significance to something is a social action, an exercise in power, which reflects the analysis by Smith (2006). The terms 'significant' and 'value' are closely related and are used to refer to a property of an object, often interchangeably. Russell and Winkworth (2009, p.1) refer to significance as *'the value and meanings that items and collections have for people and communities.'* Here it seems that significance has a wider meaning than value; it includes all the different types of value and meanings that people may apply to objects.

Attempts have been made to distinguish the types of value in the cultural heritage sector. Hutter and Throsby (2008, p.3) give a useful if brief overview of the cultural heritage literature which is concerned with perspectives on economic and cultural value. Elsewhere Throsby explains the development and theoretical background to economic theories of value (2001, p.19). He

identifies a range of cultural value characteristics, including aesthetic value; spiritual value; social value; historical value; symbolic value and authenticity and integrity (2001, p.28-29). Throsby argues that heritage goods have cultural and economic value, whereas ordinary goods only have economic value (2001, p.76). In making decisions about heritage projects, Throsby argues that a simple cost-benefit analysis does not capture all the aspects of cultural value. Much research has been done to examine the financial costs of digital preservation (see section 2.3.4). Currall and McKinney (2006) describe work by the 'eSpida' project on how to put together a business case which decision makers would understand to justify the expenditure on digital preservation. To do this they concentrate on the economic aspects of value and produce a model of value to aid this. The model seems to indicate who an information asset may be valuable to, but does not give guidance on how to measure value. It also dismisses the use of historical value, reflecting Throsby's argument that valuing heritage material economically misses other important aspects of its value. However this work does recognise that value is contingent, it is decided by the institution and the institutional context in which the decision is made will be an important. Currall and McKinney (2007, p.14) rightly state that value is often seen as an intrinsic quality whereas actually it is decided within a context.

A difficulty with value is that there is a need to operationalise a quality which is often intangible, subjective, and context-dependent. As Lunghi et al (2012, p.217) point out, 'value' is a less practical concept than 'costs' or 'benefits'. As a bridge between these concepts, the work by Tanner (2012) is interesting; he has developed the 'Balanced Value Impact Model' in which there are measures of intrinsic or intangible value, as well as measures of economic value. On examining the processes included in the model this can be most clearly seen in the first stage where the context of the resource is mapped. Different areas of impact have one or more 'value drivers' applied depending on the context (Tanner 2012, p.45). These value drivers describe the types of value that someone may gain, such as utility value where they gain benefits from being able to use the resource at some point and bequest value, where the benefit arises from knowing the material will be passed to the future. These value drivers are similar to those types of value outlined by Throsby (1997, p.16). The advantage that this model has is that it does attempt to include quantitative and monetary measures along with qualitative or intangible factors, presenting them in terms of impact.

Value can also be considered from a functional or instrumental perspective. For example Hughes (2012, p.6) lists some of the ways in which digital material may be valuable, focusing on the ability of digital material to enable things to be done, including increased or enhanced access, supporting preservation of hardcopy items, adding to the reputation of an institution and helping it to fulfil its mission. Value can be expressed through the utility of a resource or through the benefits it brings; items of high impact have high value. Tanner (2012, p.12) defines impact as:

*The measurable outcomes arising from the existence of a digital resource that demonstrate a change in the life or life opportunities of the community for which the resource is intended.*

The report for JISC by Tanner and Deegan (2011) demonstrates many benefits brought by digitised resources that could be used to measure value in a functional way, by asking does the resource bring us any such benefits. This recalls Holden's (2004) distinction between intrinsic or instrumental value; the latter is measured in its usefulness to government policy programmes (2004, p.21). Charles Beagrie Ltd (2012) has created a toolkit for assessing the benefits of long term preservation of digital research data in higher education, based on the 'Keeping Research Data Safe'(KRDS and KRDS2) projects. A similar toolkit for measuring benefits and impact of digitised scholarly resources was created by Meyer et al (2009), for JISC; methods of assessing impact investigated were both quantitative and qualitative, and the toolkit describes the different methodologies and gives case studies as examples. Beagrie et al (2011, p.56) point out that the outcome of the preservation process that is important is not the set of preserved information, but the value of the activities performed with the preserved material. An example can be seen in research for the HLF by Flow Associates and The Collections Trust (2010, p.12), where respondents were asked what they thought digital heritage is, and their answers included examples of activities that constitute or exemplify digital heritage. The emphasis was on activity and use, such as delivering content and services or using digital tools to increase user interaction and access.

The literature more often emphasises the differences between selection and appraisal in each domain; each has its own body of literature as discussed previously relating to traditional forms of selection for traditional material. However in the digital context some authors, such as Harvey (2007, p.27) and the DPC Decision Tree (2006), focus value and criteria suggestions on the material more than on the institutional context. Harvey (2007, p.27) for example, whilst earlier recognising different types of value used in archives and libraries, goes on to suggest that a key question is 'why are the materials worth keeping?', associating value with evidence, information, artistic or aesthetic factors, significant innovation, historic or cultural associations, users and cultural significance. This overall approach has advantages in enabling the sharing of information between domains, de-emphasising differences. The authors of the Significance 2.0 framework argue that:

*Nevertheless, in a collections environment that is increasingly connected and intertwined, there are benefits in using a flexible process and a common set of assessment criteria to share the meaning of collections and explain how and why they are of value.*

Criteria are necessary to provide guidance on selection but there is no agreement on which are most appropriate or which types of value are most important. This will depend on the context in

which the institution operates; the value of material will vary between institutions and stakeholders (see section 2.3.1). Archives, libraries and museums will emphasise different types of value which relate to their purpose. Value is not a static characteristic. It may change over time - an academic library could find some of its collection is of less value when the university subject priorities change for example. As UNESCO (2003, p.72) warn, there are dangers in assuming that current value assessments are a reliable guide to future values. Thompson (1979) proposed a 'theory of rubbish', in which value moves between three categories: *transient* which diminishes over the life of the object and turns to *rubbish*, which in some cases may increase in value and become *durable*. So objects with transient value are produced, most become rubbish but some of the rubbish category then become durable, so in time rubbish can become durable. Whilst this can clearly be seen with tangible heritage or hardcopy documents it is difficult to apply this to digital material as there is little opportunity for digital material to spend any time being rubbish, due to the need for it to be curated early in its lifecycle.

However value is perceived, the literature reveals that it is a social construct (Lloyd 2007, p.54; Cameron 2007, p.57) and selection for digital preservation is a social and cultural process (Lavoie and Dempsey 2004). The use of value as a criterion in selection is not unproblematic; material has different value to different stakeholders, leading to the question of whose definition of value to use. Marchionni (2009) argues that users partly determine the value of digital material, from when it is created, due to their use of it, although this view describes an indirect effect that still requires interpretation by practitioners. The question of user involvement in determining value and selecting material relates to debates in the wider heritage management literature. Cameron (2008, p.180) and Smith (2006, p.12) questioned the hegemony of heritage management and highlighted the role of non-experts in using and creating heritage in their own way. Whilst the idea of including users in selection is a worthy one, it seems difficult to see how this could be achieved in a practical way and whether users would want to be involved. However some projects have been conducted in which users have been involved in 'crowd sourcing' activities, which include users in the creation and gathering of material, such as in the British Library's UK Soundmap (Pennock and Clark 2011), or the World War 2 'Peoples War' project by the BBC (2006). What is clear from the literature is that whilst there is acknowledgment of the potential role of users in selection of material, there are very few examples of this occurring and then only in a prescribed manner, and the role of memory institutions is to act on their behalf. The next section will examine roles and responsibilities more closely.

### **2.3.3 Roles and responsibilities**

Digital material has multiple stakeholders: creators, publishers, rights holders, librarians and users for example. There may be a wide range of stakeholders that have responsibility for digital material at some point in its lifecycle; different stakeholders have influence on and

interest in preservation at different stages (Lavoie and Dempsey 2004; DPC 2008, p.65). The literature reflects the assumption that part of the role of libraries, archives and museums is preservation of cultural memory and this continues in the digital context (Feather 1996, p.58; Usherwood, Wilson and Bryman 2005). However the assumption that libraries, archives and museums are best placed to select which items of digital cultural heritage are to be preserved has been challenged; Gladney (2008) and Bearman (2007, p.33) assert that the institution-centred model of preservation is inefficient and will fail. Bearman suggests moving the emphasis from the institution to the systemic level, in which preservation takes place on the network (2007, p.35) so only a few copies need to be kept. Whilst this view is thought provoking, it seems currently unrealistic as there are many barriers such as issues of trust and copyright (see section 2.3.5) to overcome.

The commitment and resources necessary to preserve material especially in the long term is very great and so the acceptance of responsibility has wide implications for an institution. This is made more difficult as the length of time this responsibility is for is often unclear; there is ambiguity found in the digital preservation literature relating to time scales. A report from PORTICO (2011) refines the definition of digital preservation though it is not made clear what 'very long term' may mean:

*Digital preservation is the series of management policies and activities necessary to ensure the enduring usability, authenticity, discoverability and accessibility of content over the very long term. (Portico 2011, p.27)*

The lack of clarity in determining time frames for digital preservation is reflected in the definition given of 'long term' in the OAIS reference model:

*'...long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community. Long Term may extend indefinitely. (CCSDS 2012, p.1-1)*

There is an implication of 'in perpetuity' which then has further implications for resources and whether an institution that may be willing or able to take on this responsibility. Incentives to take part in digital preservation have been explored by Lavoie and others (for example, Lavoie 2003; Currall and McKinney 2006; BRTF 2010). Currall and McKinney (2006) argue that as information and its value is intangible, it is difficult for organisations to understand the benefits of investment in curation, so they do not have the incentive to preserve material. The DPC Handbook (2008) recognises that:

*an institution may be reluctant to take primary preservation responsibility for materials it acquires if it feels that interest in its preservation is so widely shared that it would constitute an unfair burden on their own institution.*

Despite this it seems that institutions are recognising responsibility for preservation; the APARSEN (Alliance for Permanent Access to the Records of Science Network) project report examining business preparedness in research libraries found that:

*Overall, 70% of organisations have developed specific objectives pertaining DP [digital preservation], 59% have incorporated DP into their strategic plan and 55% have incorporated it into their mission and vision statements (Riestra et al 2013, p.22).*

Another theme found in the literature relates to the question of shared responsibility. Jones and Semple (2006) describe how responsibility within an institution for preservation is unclear as it could be seen as 'an IT problem', neglecting the valuable input of information and curatorial professionals, especially in selection of material. In order to manage digital material there may be a need for librarians, archivists and curators to share responsibility with IT staff; for example Seadle et al (2012, p.171) describe in their case study of the UK LOCKSS (Lots of Copies Keep Stuff Safe) Alliance that '*Staff responsibilities tend to be split between librarians responsible for collection development and IT staff responsible for system maintenance*'. On a national level, Verheul (2006, p. 29) in her survey of fifteen national libraries found that whilst the libraries all had at least one unit or department that referred to digital objects in some way, none of the libraries had placed all digital preservation activities in one unit. She found that co-operation between departments was often through formalised arrangements with cross-domain working groups (2006, p.31). As Jones and Semple (2006) rightly point out, digital preservation needs to be a cross-disciplinary responsibility as relevant skills may be spread throughout an organisation. A similar conclusion was reached by Runardotter et al (2011, p.76) in an examination of responsibility in archives; they found that '*cooperation, coordination and communication*' between archivists, IT personnel and managers with strategic responsibilities is needed to underpin shared responsibility for digital preservation. They note that these are often missing and archivists alone are responsible for digital preservation, whilst their influence is limited. The relationship of archivists and record managers with IT staff has been explored on a small scale by Oliver, Chawner and Lui (2011) who concluded that firstly archivists and IT people are not working together, and secondly that IT staff view record keepers as having expertise only with paper records and there is a lack of confidence in their ability to manage digital records; the cultural differences between IT people and record managers were very wide (2011, p.321). A lack of common understanding and ability to work together could jeopardise organisational efforts to preserve digital material and a co-ordinated approach is necessary. From early in the digital preservation literature the importance of clear responsibility for decision making and preservation has been recognised (Waters and Garrett 1996; Haynes et al 1997;

Eden 1997; Ayris 1998). Few individuals that become involved with the development or management of digital resources have influence over those resources throughout their entire lifecycle. This can lead to a lack of clarity regarding responsibility, both within and between institutions. The 'Mind the Gap' report (Waller and Sharpe 2006, p.16) found only 33% of organisations surveyed had a clear responsibility structure for digital preservation.

Digital preservation is not only an institutional responsibility but also a cross-domain issue in which memory institutions from different domains have an interest. Digital objects are not dependent on a geographical location, as shown by the development of cloud computing services; there is less need to visit the institution (Feather 2006, p.12). Also as Dempsey (1999) comments:

*The user wants resources bundled in terms of their own interests and needs, not determined by the constraints of media, the capabilities of the supplier, or by arbitrary historical practices.*

These factors may act as drivers toward shared services, practices and policies for selection and preservation between traditionally separate domains. Collaboration is useful in digitisation projects; a survey by Portico (2011, p.15) reports that 63% of the institutions they surveyed were collaborating with other departments within their own organisation and 57% were participating in intra-institutional collaborations when digitising books. Practitioners are able to source the knowledge and expertise from other stakeholders when managing digital material. Collaboration also aids in spreading the costs of elements of digital preservation including infrastructure and training of staff (Baker and Evans 2009, p.12; Lunghi et al 2012, p.214). Many examples of organisations enabling collaborative approaches to digital preservation are described by Angevaere et al (2012, p.95). Zorich, Waibel and Erway (2008) examined collaboration between libraries, archives and museums, recognising that the concept includes many different types of co-operation from weaker contact to full convergence of services and identity. Their report only focused on libraries, archives and museums within organisations; an example of collaboration on a national scale can be seen in Canada where the National Archives and the National Library merged to become Libraries and Archives Canada (LAC) in 2004. The combined institution has a single digital collection policy and digital preservation policy. The issue of convergence in a digital context has been examined by Katre (2011), Trant (2009) and Given and McTavish (2010, p.22). All identify separate educational regimes for librarians, archivists and curators as a factor which hinders different domains working together to managing digital material as it emphasises differences rather than similarities (Trant 2009, p.376).

A particular problem for many libraries which illustrates the lack of clarity of responsibilities is that of preserving e-journals and other electronic resources to which a library leases access through licences from publishers. With hard-copy material if a journal subscription was cancelled the library would still have ownership, access and responsibility for the back issues

they had bought. However this is not always the case with licensed digital material, though libraries are able to arrange or purchase access to back files; the NESLI2 model licence negotiated by JISC for example includes clauses relating to continued access after the termination of the licence (JISC n.d.a). There is also the LOCKSS (Lots of Copies Keep Stuff Safe) service developed by Stanford University in the USA, which is a collaborative service that stores copies of e-journals from publishers (with their permission) so libraries are able to collect, store e-journal material and continue to access it even if the publisher went out of business or their subscription ended (LOCKSS n.d.). Responsibility for preserving licensed material such as e-journals is a complex issue and it is unclear where this should lie; the factors which influence the choice of approach to e-journal preservation are outlined in the JISC report into e-journal archiving solutions (Morrow et al 2008), and this includes changing responsibilities (Morrow et al 2008, p.9-10). Interested parties include publishers, libraries, institutional repositories (where an institutions research output is collected, preserved and made available), legal deposit libraries, and third party services such as Portico. It is clear that unless there are specific arrangements for archiving the material, such as archival clauses in licences, preservation and on-going access to the material by libraries cannot be guaranteed. In response to this, some libraries have taken responsibility for preserving their licensed material. Muir (2004, p.80-81) found in her survey of libraries that the majority already took responsibility for preserving their digital collection and this included 17.3% which took responsibility for preserving material to which they only had licensed access. More recent surveys have been performed on this issue amongst different groups of stakeholders, such as those by the PARSE.Insight project (Kuipers and van der Hoeven 2009) who surveyed researchers, data managers and publishers from the EU, USA and elsewhere; Durrant (2008), who surveyed publishers and Meddings (2011), who conducted a survey of academic libraries. These all report a desire for national libraries to have a large role in preserving e-journal material. For example, Durrant found that 70% of respondents from the Association of Learned and Professional Society Publishers saw the responsibility for long term preservation of electronic scholarly material to lie with national or legal deposit libraries (Durrant 2008, p.6). Only 44% felt each publisher could create its own repository for its digital content. Sustainable preservation needs responsibility taking not only for the selection and preservation of material but also for funding and it is unlikely that only one set of stakeholders will be able to do this.

Much of the literature focuses on the roles and responsibilities of the institution, but selection is performed by the individual, so it is necessary to consider their role and responsibilities also. As already discussed, there may be difficulties where the role of the practitioner intersects the role of IT staff (Jones and Semple 2006; Runardotter et al 2011; Oliver, Chawner and Lui 2011). Seadle (2012) conducted an ethnographic examination of the LISA (Library and Information Science Abstracts) database, which includes much, though not all, of the library orientated literature. Despite the somewhat ad-hoc nature of his investigation, he finds that the focus of articles contained in LISA is more on repositories and metadata, rather than technical aspects

(2012, p.375). He concludes that the community represented in LISA is concerned about the long-term future, but has not come to terms with the necessary technical issues. However this assumes that those who read this type of material need to have a great deal of technical knowledge - is this part of the role of librarians? Conway (2010, p.73) presents the need for expertise in technical as well as curatorial skills as a dilemma, between training existing staff and keeping the responsibility for digital preservation in house, or to take the risk of outsourcing it to third parties that already have the expertise, but may view the activities from a different perspective. He attributes this dilemma to the gap between the immediate need for expertise and the time it takes to train or recruit appropriate staff. The dilemma is not as clear as presented by Conway as the whole process need not be outsourced; or the institution can enter in to specific contractual agreements with third parties, which protects their control and curatorial influence.

Craig (2004, p.75) identifies some of the potential difficulties encountered by archivists when faced with new forms of technology. These include a rapid transformation from established procedures to new systems; inadequate change management planning; and technical stakeholders that do not understand the archival point of view. These factors are dealt with only briefly and the response of practitioners to the introduction of new technology is under examined in the wider library and museum sectors also. Many reports have identified a strong need for digital preservation training (for example the DPC Training needs analysis report (2004); Simpson (2005, p.22); Waller and Sharpe (2006, p.31); RLUK (Research Libraries UK) and the BL (Arthur 2009)). Training and education in digital preservation is available; the DigCurV website has compiled a wide list of digital preservation training courses, although many are not in the UK. In the UK introductory courses are run by organisations such as the DPC, the British Library Preservation Advisory Centre, the APARSEN network and JISC. An example is the Digital Preservation Training Programme (DTPT) from the DPC and the University of London Computer Centre (ULCC). In addition there are Masters level courses available through a few UK universities which either focus on or have a large proportion of their content which focuses on digital preservation, including Aberystwyth, Dundee and Northumbria. However not all those who are managing digital material have been on university courses which include a digital element, nor have been able to attend training courses. The DTPT course for instance only runs twice a year (DTPT n.d.). The DPC Handbook (DPC 2008) recognises that management support is needed in order to train people adequately. This in part needs to be 'hands on' but courses may not provide this; the DTPT course does not provide 'hands on' experience with tools for example.

#### **2.3.4 Finance and costs**

The argument is pervasive throughout the literature that there are inadequate resources to preserve all digital information so selection is necessary (for example Russell 1999; Harvey 2007, p.8; Berger 2009, p.60), although this view has been challenged (Rusbridge 2006). Whilst

the issue of resources is not limited to digital material, it becomes pressing when the cost of digital storage, expertise in information technology, curatorial skills and the need for constant interventions to keep material useable is considered. Factors such as discovery become harder with increased volumes of material and preservation metadata describing context and content need to be created to aid discoverability, manage resources and facilitate reuse (Day 2005, p.12); the time and expertise to create the necessary metadata is an additional factor in the overall cost to preserve. A second and related reason is that the volume of digital information that is available to be preserved is too large for it to be feasible to preserve everything, so selection is inevitable to identify material of enduring value (for example Feeney 1999, p.11; Harvey 2007, p.9; Deegan and Tanner 2006, p.15; Ooghe and Moreels 2009). We cannot save everything. This assumption has been challenged in the literature, particularly on the basis of cost. Bearman suggested if preservation takes place on the network, everything could be selected due to economies of scale (2007, p.35) and Bailey (2008, p.100) argued that it would be possible in a record management context to keep everything as storage is huge and cheap and search tools could be employed so there would be no need to manage the information. Neumayer and Rauber (2007) in their deliberately provocative position paper point out that appraisal in its current form is very expensive because of the costs of trained staff and equipment; they suggest a form of random selection to be more cost effective. This strategy would also have the advantages of scalability, be fair and unbiased and be simple. The literature on costs and selection often focus on the costs that can be saved by selection; the question is asked 'does the cost of selection outweigh the cost of preservation activities?' and the answer given is usually 'no' (Whyte and Wilson 2010; PARADIGM 2008). This is attributed to the much higher costs of creating or providing adequate metadata which is necessary to preserve material, the large quantity of material and the implications for future costs in taking responsibility for material; selecting material is seen as a cheaper alternative.

Sustainability is central to digital preservation in anything but the very short term; economic sustainability for preservation is defined as a *'...means of keeping information accessible and usable over time by ensuring the ongoing and efficient allocation of resources to its maintenance'* (BRTF 2010, p. 107). Law (2010, p.37) points out that the overall costs of traditional libraries are not well understood, although budgets are, and the true costs of digital libraries are also unclear. The costs to set up a preservation service within a library or archive are great and there is difficulty in predicting future costs which may not be provided for initially (Sierman 2010, p.118). Law (2010, p.38) does note that the change to a digital from a traditional library may include savings from a reduction in physical space and that some of the costs will move from the library to the user, such as equipment. Whilst the cost of digital storage is likely to reduce, the ever increasing volume, and the development of different formats and media of material which could then be included in the remit of a library or archive is unpredictable (Kinnaes et al 2010, p.97). Whilst funding is forthcoming for creating digital resources, there is as yet no solution to the problem of funding long term preservation of those resources,

especially when many digital preservation initiatives and activities are supported on a project or short term financial basis (Maron, Yun and Pickle 2013, p.11). The CARARE project, part of the Europeana Project Group (Moore, Jeffery and Richards 2010) takes a broad view, reviewing different funding and income generation models for those cultural heritage organisations with a role in generating, curating and disseminating digital cultural heritage. The types of funding identified include government funding, project funding, endowments, commercial partnerships, direct charging, collaboration and from multiple streams. The authors conclude that ‘*no single funding model can currently be considered as 100% sustainable*’. Whilst the benefits of multiple funding streams in order to spread the risk (Baker and Evans 2009, p.11) are clear through being more robust than single approaches, the CARARE authors note that collaborations and mixed funding models have additional complexity from extra administration (Moore, Jeffery and Richards 2010, p.53). Various projects have examined the issue of lifecycle costs and business models for economically sustainable curation and preservation and Lunghi et al (2012, p.199) give a succinct overview of the findings from these and other research into costs. The LIFE series of projects (McLeod, Wheatley and Ayriss 2006; Ayriss et al 2008; Hole et al 2010) models the lifecycle costs of the activities involved in curating digital material, each iteration of the project presenting a more refined version of their lifecycle costs model, and it includes selection as part of an acquisition phase (Hole et al 2010, p.84). The LIFE3 project produced a web-based predictive costing tool (Hole et al 2010, p.89). Further work has been undertaken in the KRDS (Keeping Research Data Safe) (Beagrie et al 2008) and KRDS2 projects (Beagrie, Lavoie and Woollard 2010), which build on the LIFE work and other relevant initiatives (Beagrie, Lavoie and Woollard 2010, p.8) to develop the KRDS2 Activity Model, a lifecycle costing method applied to research data. The framework details the key variables that may have an effect in cost, the activities that need to be assessed, and a template to help institutions model their costs (Beagrie, Lavoie and Woollard 2010, p.11). Selection is included in a small way as part of the acquisition activity (2010, p.17). Conclusions from KRDS2 include that the cost of archiving is very small in comparison to the overall costs, especially access and acquisition/ingest and that potential cost benefits may come from the development of tools to aid ingest and access (2010, p.79).

With limited resources and unclear business models, selection of digital material for preservation, or more specifically for preservation actions, becomes even more important. As the BRTF report (2010, p.46) makes clear, when considering the on-going costs of preservation and the potential for improving cost-efficiency, selection criteria to prioritise preservation investment are critical to sustainable preservation. The report makes selection of materials with long term value one of its five conditions to achieve economic sustainability; selection is necessary to give priority to ‘*...materials that have the greatest promise of returning value to users over time.*’ (BRTF 2010, p.76). Careful selection to target resources at preserving the most valuable material is essential to provide a sustainable service. The BRTF (2010, p.37) ask the question that as there is no way of predicting what future users may need, on what basis

could institutions select items now? They recommend an 'option strategy' as a potential way forward. They reiterate the problem of non-selection equating to deletion and the cost this could involve as the decision is irreversible. Instead they suggest that a small investment in 'holding' the material may be justified, in case it becomes important later. Decision makers would be 'purchasing an option' to put off final decision making. This seems an attractive alternative to taking immediate responsibility for something that may not be useful in the future, but there is no way to predict when demand may increase and it would still require funding and expertise to preserve access to the material until a decision can be made. Regardless of the strategy adopted, the selection of digital material in order to help control costs of digital preservation should have a high priority within institutions, which may require a change in strategic priorities (Lunghi et al 2012, p.220).

### **2.3.5 Legal issues, ethics and risk**

The context in which any institution operates is bound by a legal and ethical framework, which is also related to managing risks to the institution. Vermaaten, Lavoie and Caplan (2012) discuss an approach to risk assessment for digital preservation, giving an overview and comparison of previous work and typologies of risks. They point out that risks relating to digital preservation are of two types. Firstly there is risk to the material itself from technical issues, and secondly there are risks resulting from the wider organisational context. Clifton (2005, p.21) lists risks to digital material and includes collection level risks, such as having appropriate metadata, restrictions from intellectual property rights and a lack of knowledge about digital holdings in the collection. In addition, he identifies organisational risk factors, including a lack of appropriate policies and resources. In order to assess organisational risks, the DCC and Digital Preservation Europe (DPE) developed DRAMBORA (Digital Repository Audit Method Based On Risk Assessment), a methodology for the self-assessment of risks in repositories, presented as an online tool. DRAMBORA helps managers to map the organisation in order to identify areas of risk and quantify them, as a way of auditing the institution's risks and strategies. It includes staff, the mission of the repository and activities (Donnelly et al 2009, p.4) rather than only focusing on technical risks to the material. In addition the TRAC (Trusted Repositories Audit and Certification) Criteria and Checklist (RLG-NARA Digital Repository Certification Task Force 2007) may aid institutions to identify and understand the risks to their repository (RLG-NARA Digital Repository Certification Task Force 2007, p.3). There is a generic Risk Management Standard (ISO/FDIS 31000, 2009) which could be used to underpin approaches to risk management (Barateiro et al 2010). In addition to risks to the material, there are also risks to the organisation from not adhering to legal requirements; this has an effect on the material selected for preservation, such as where an organisation chooses to keep records of activities, and how it is subsequently managed. Risks could arise from issues such as allowing access to potentially sensitive or offensive material or not owning rights to copy or change the material (Muir, Buttler and Mossink 2012, p.66).

Copyright and related rights are important factors in digital preservation activities, including selection. Practitioners wishing to make copies of material for preservation purposes, for instance in order to move content to new storage media when the old media becomes obsolete, may need to obtain the copyright owners' permission; Muir (2004, p.76) outlines the rights implications of various preservation strategies. The relevant law varies internationally. Muir's position piece (2006) aimed to provoke debate around the potential impact of current legal provisions on the preservation of digital materials in libraries in the UK and Besek et al (2008) provide an overview of the impact of law on digital preservation in Australia, the Netherlands, the UK and the USA. The KEEP (Keeping Emulation Environments Portable) project has produced a 'Layman's guide' (Anderson 2011) to the relevant European Community law and international treaties that affect the work of the project and the legal position in the Netherlands, Germany and France. In the UK, the relevant act is the Copyright, Designs and Patent Act 1988 (Great Britain 1988), as amended. The Act includes a preservation exception (s. 42) whereby prescribed archives and libraries are allowed to make one copy for preservation purposes, providing certain restrictive conditions are met. These include that the material is for reference only and that it is a literary, dramatic or musical work, not artistic or a sound recording. The Gowers Review of Intellectual Property (2006) recognised the difficulties faced by libraries and other institutions that are unable to legally make copies of works for preservation purposes. Recommendations 10a and 10b from this report refer to allowing institutions to make copies of all types of material in order to alleviate wear and tear and to shift formats to avoid obsolescence (Gowers 2006, p.66). The Intellectual Property Office (IPO) has worked to take the recommendations from the Gowers review forward, with a two stage consultation from 2008, ending in 2010. The difficulties faced by institutions in preserving content was yet again recognised by the Hargreaves report (2011, p. 50).

National libraries have a remit to collect documentary heritage of national interest regardless of format. The UK Legal Deposit Libraries Act 2003 required secondary legislation to expand the scope of legal deposit beyond print material and to bring a preservation exception for legal deposit libraries (s. 44A of the Copyright, Designs and Patents Act 1988) into effect. A voluntary code for publishers to deposit microfilm and offline digital material with the British Library, one of six legal deposit libraries in the UK and the Republic of Ireland, has been in place since 2000 and updated in 2010. A pilot scheme for scholarly electronic journals was begun in 2007 (British Library, n.d). The Legal Deposit Advisory Panel (LDAP) was established to understand the conceptual and practical problems of extending legal deposit (Gibby and Green 2008) and to make recommendations on further regulation. The Department for Culture Media and Sport launched a public consultation exercise in December 2009 regarding the panel's recommendations and a further consultation, including draft regulations, in 2010. LDAP was disbanded in 2010 by the then new government in a review of public bodies (DCMS 2010). It is clear from the government response to the last consultation (DCMS 2011) that both libraries and publishers had numerous concerns about potential regulations and there were many issues

to be agreed upon between the various stakeholders before any regulations could be introduced. The Legal Deposit Libraries (Non-Print Works) Regulations 2013 were laid before Parliament at the end of January 2013, and came into effect 6th April 2013. As well as enabling the deposit of non-print material (s.15 - s.18), these regulations allow the deposit library to copy the material for preservation purposes (s.29).

Ethical issues are relevant in selection; selection, appraisal or acquisition is referred to in the codes of conduct or ethics for librarians, archivists and curators devised by international bodies, domain specific professional organisations. Berger (2009) examined the role of ethics in decision making for digital preservation. In her view preservation is a by-product of the ethical imperative to provide access to material as items need to be preserved to have access to them for longer. But as we are unable to keep everything, the ethics of preservation becomes closely related to the issue of selection (2009, p. 60). A further ethical issue arises when the creation of criteria for selection is considered. Lloyd (2007, p. 60), argues that:

*The development of criteria... underlies the subjective positions and political interests of those charged with determining significance and thus privileges some memories over others.*

Lloyd goes on to argue (2007, p.63):

*Assessing an item's value... against a formulated set of criteria appears reductionist; it assumes that core values and beliefs about what is worth remembering are common to the diverse groups that constitute a society.*

The question of whose values are being used in the formulation of criteria does not seem to be answered within the digital preservation literature. This recalls debates in the wider heritage literature, in which the disenfranchisement of minority voices in selecting heritage is clear.

### **2.3.6 Policy**

Selection should be driven by decision making and policy rather than be dictated by chance, time or technology. '*Without policies, practitioners have little to guide their decisions about what must, should, could and won't be preserved, let alone how*' (Emmott 2008). The definition of a policy used here is:

*a formal statement of direction or guidance as to how an organization will carry out its mandate, functions or activities, motivated by determined interests or program* (Interpares2, 2008b).

In order to make apparent the aims and methods of selection within an institution it is recommended in the literature that it develops clear policies. The UNESCO guidelines for the preservation of digital heritage (2003, p.59) state reasons for an organisation to create a preservation policy:

*Preservation programmes should be guided by a policy framework that says what the programme is trying to do and how it will try to achieve it. In a field of such complexity and evolving understandings, a policy document needs to provide clear, long-term direction as well as regularly reviewed guidance.*

A policy is necessary to provide direction and guidance to an internal audience and to define why an organisation is doing digital preservation, for both internal and external stakeholders. Funding and strategy decisions may need to be justified. A policy also acts as the authority for those undertaking digital preservation (TNA 2011b, p.5). It may be appropriate and useful for an organisation to create an aspirational policy if a working policy is not yet appropriate; this makes a statement of commitment to digital preservation even if specific activities are still being developed (TNA 2011b, p.6); such a policy may have reputational value to an organisation.

A preservation policy is critical to libraries as preservation activities need to be considered within the overall collection management approach (CEDARS 2002, p.12-13). In addition the requirement for a policy can be imposed; UK research funders require institutions to have a number of policies relating to areas such as data creation, metadata, appraisal, storage and preservation (Jones 2009, p.26). Jones (2009, p.27) examines the provision of digital preservation policies in institutions, though unfortunately this part of the report is brief. She finds that policies for projects and for repositories are most common, whereas institutional level policies are lacking. Research has consistently found a lack of digital preservation policies within institutions (for example Ayre and Muir 2004, p.107; Waller and Sharpe 2006, p.16; Beagrie, Rettberg and Williams 2008, p.1). Beagrie, Rettberg and Williams (2008, p.1) highlight the lack of digital preservation policies in libraries. Their work was funded by JISC so focuses on the higher education domain in particular, but they point out that there is a similar picture across many domains. A survey for the Planets project (Preservation and Long-term Access through Networked Services) found that 43% of libraries who responded had a digital preservation policy (Sinclair et al 2009, p.274). While this figure may seem low, it demonstrates an increase since the survey in 2004 by Ayre and Muir (2004, p.107) who found that only four out of sixty nine libraries who responded had digital preservation policies. Unsurprisingly, Sinclair et al (2009, p. 280) also found that those organisations with a policy were much more likely to have solutions, budgets and plans in place for the long term management of digital material. What these surveys do not uncover though are the barriers to libraries and other institutions engaging sufficiently with the need for policy development.

Whilst the importance of having a selection policy for digital preservation is recognised (UNESCO 2003, p.71 for example) there is little guidance specifically on creating selection policies for digital preservation. Whyte and Wilson (2010) have produced guidance on selection for the curation of research data in which they discuss selection policies, although this section is brief. Beagrie et al (2008, p.20) recommend including in a digital preservation policy a section on 'identification of content', which details the material to which the policy applies, including its relative priority or value. This clause could include appropriate selection principles. Of particular interest here is the DPC Decision Tree (2006) which is explicitly concerned with providing guidance on creating selection policies. It takes the form of a decision map of questions relating to selection, rights and responsibilities, technical and cost issues, and documentation and metadata. The questions are highly relevant to policies but there seems little on user demand, except in terms of which format they might prefer, or perhaps in a reference to 'long term value'. As this is a generic guide the meaning of 'long term value' is left undefined. The section focussed on technical issues of storage and file formats also includes costs, which then removes the issue of costs from consideration in conjunction with other issues, such as value, use and acquisition. Whilst the format of the Decision Tree is helpful as earlier questions should be answered before later ones, it is not easy to use as a guide to creating a policy. It neglects to mention inclusion of clauses relating to the policy itself, such as responsibility for the policy or when a review will take place. In addition many of the questions refer to further guides or model policies which are no longer available. It is clear that the guide, whilst useful in highlighting many of the issues which need to be taken account of when considering selection, needs updating and revising.

Although each institution has its own priorities, the lack of consistent terminology within guides and publically available policies, which can be used as exemplars, is problematic for those seeking to formulate policies of their own. The definitions of 'policy' and 'strategy' are not consistent (Dappert et al 2008, p.9). However there is guidance available to cultural heritage institutions who are considering formulating a policy for digital preservation, including the ERPANET policy tool (2003) and TNA guidance (2011). Beagrie et al (2008) have developed an outline model preservation policy, based on research for JISC, to guide institutions in formulating a digital preservation policy. The National Preservation Office at the British Library, now the Preservation Advisory Centre, also published a short booklet giving guidance on preservation policies which briefly mentions digital material (Foot 2001). Despite the availability of policy guidance there is no agreement within them on what preservation policies should contain, reflecting the different purposes and domains in which these policy guides are to be used.

The lack of guidance for selection policies has driven an examination in this research of the wider category of digital preservation policy guides, which may contain recommendations for clauses or provisions relating to selection. The digital preservation literature was searched for

documents and articles giving advice on how to create a preservation policy. Beagrie et al (2008) was chosen as a convenient starting point for the analysis as it is particularly comprehensive and detailed, basing its findings on an analysis of existing institutional policies from different contexts. Appendix 1 details the explanations given by Beagrie et al (2008) for each of the clauses they recommend. The definitions given by Beagrie et al (2008) for the suggested clauses are broad, in line with the function of the recommendations and the need to be inclusive of many different institutional needs, but they are ultimately confused. Each policy clause has very many potential sections, many of which overlap. Some of the clauses suggested conflate description and management of the policy with description and management of the activities of the institution. A single clause can describe what the policy should say about itself and then also be concerned with similar actions in the institution. For example, it is suggested that the section on preservation objectives includes why the policy has been created, alongside 'principle aspects for implementation' and 'high level preservation guarantees'. The preservation strategies section includes consideration of the legal environment, as does the intellectual property section. Analysis of the policy guides shows that there is a need to clearly separate the clauses relating to the continued functioning of the policy from the clauses relating to the material the policy affects. For example, having a clear section relating to the responsibility for and review period of the policy would improve the guide. These issues have made analysis of the recommendations more difficult. It also means that a policy guide has been counted as addressing each of these clauses if it has mentioned some, but not necessarily all, the items in the clause as listed by Beagrie et al (2008).

In addition to Beagrie et al (2008), the ERPANET policy tool (2003) and the guidance from TNA (2011b) were examined, along with:

- ERPANET Digital Preservation Policy tool (2003). ERPANET was a European Commission funded project which aimed to establish best practice and skills development in digital preservation of cultural heritage and scientific objects'. It ran from 2002 to 2004.
- CHIN (2006) refers to an online tutorial from the Canadian Heritage Information Network called Concepts for Developing Digital Preservation Policies, aimed at museums in particular
- ICPSR model (McGovern 2007) is described as providing '*an outline for constructing the digital preservation policy framework for ICPSR and offers a step towards identifying core components of a digital preservation policy framework to encourage a community standard for digital preservation policy documents*'. It originates from the University of Michigan, USA.
- JISC 'Establishing a Digital Preservation Policy'(2009) paper
- DCC Preservation Policy Template (Jones 2010) aimed at repositories

Table 18 in appendix 2 describes the originating author, function, and audience in order to help clarify the contextual differences between the guides. There are broadly two types of preservation policy assumed in the guides, high level strategic documents and practical 'how-to' documents. Beagrie et al (2008, p.16) suggests one policy could have both types of information or that these could be in different documents. Comparison between the guides is not always straightforward as the terminology is inconsistent; for example the term 'preservation strategy' may be used to mean the technological solutions to digital preservation (similar to Xie 2007, p.2), reflecting the point made by Dappert et al (2008, p.9) that the terms strategy and policy are not used consistently. Recommended clauses, definitions and the descriptions used differ by function and audience of the guide, and context in which it has been made. For example, the TNA guide is concerned with archives which accept records, so it is not concerned with content selection but is with the process for accessing records. The ISPCR model and Beagrie et al (2008) are most similar, reflecting the similar higher education, research intensive, institutional contexts on which they are focused. By examining the guides it is possible to identify common or 'core' clauses. Table 19 in appendix 3 shows the comparison matrix for the policy guidance documents. Clauses recommended by all, or all but one, of the guides include the following:

- Principle statement: defined by Beagrie et al (2008, p.16) as '*address how the digital preservation policy can serve the needs of the organisation and the benefits it will bring*'.
- Preservation objectives: includes why the policy has been created; any high level guarantees or exclusions as to what will be preserved; commitment to risk assessments and auditing; any general (legal) policies that will be adhered to (Beagrie et al 2008, p.19-20)
- Preservation strategies: give detail on the technical approach taken to ingest, storage, data management, administration and access (Beagrie et al 2008, p.32-33)
- Contextual links with other policies or documents
- Identification of content
- Standards compliance

These clauses set the scope of collecting and put it into the context of other activities. These clauses speak to both internal audiences that need guidance on the scope of their collecting activities and the relevant standards which they may need to be aware of, and to external audiences by demonstrating that the institution is aware of relevant standards and also what it will not collect. The earliest guide from the ERPANET project recommends including in the purpose of the policy consideration of the value of the material (2003, p.6) which is not present in subsequent guides. The suggested clause that applies most directly to the question of selection is 'identification of content', described by Beagrie et al (2008, p.20) as including the following:

*List a high level overview of what materials are to be preserved.....Once each category is identified, state how long each one is to be preserved and how to access it. If necessary, state what is definitely not preserved, for example certain file formats.*

All the guides apart from the ERPANET guide include 'identification of content' as a recommended clause, though often as a 'high level' list, which could be very broad. It is clear that the guides recommend inclusion of selection as a policy clause and this will be examined further by comparing actual policies to the guides to see if this recommendation is followed (see section 5.5.5).

### **2.3.7 Summary**

This section has explored the digital preservation literature to discuss aspects of selection. Selection is addressed often tangentially or in terms of practical criteria. The influential OAIS reference model ignores what is termed 'pre-ingest' (Beedham et al 2005, p.26), where selection could take place. Criteria used in memory institutions are varied, though 'use', technical ability, legal rights and permissions, documentation and value are common criteria found in the literature. There is a clear need to prioritise preservation of the most valued material though this is based on the assumption that agreement on what this is can be reached. It is clear that the concept of value is multifaceted; different types of value such as economic value, cultural value, heritage value and aesthetic value are all important, depending on the context of the valuation. Terms used in the literature, especially 'significance', bring added complexity. Ethical questions of how criteria are developed have been asked and professional codes of practice exist to guide collecting. In addition, legal constraints especially from copyright restrictions may affect selection as might the requirement to manage institutional risks.

Despite attempts to devise overarching guides, the different roles and responsibilities that institutions traditionally have influence their selection behaviour and the assumption that it is the role of the practitioner to select is still prevalent. It is still firmly within the role of the professional, who has specific knowledge and skills to allow them to do this effectively. Users and selectors may not value items in the same way, but users and creators are subject to judgements by practitioners. There are drivers toward a shared responsibility both within institutions and between different domains. A more pertinent question might be 'who pays?' and other stakeholders, especially funders, may have an influence on selection. There is no single appropriate funding model. Selection itself is expensive but is seen as a cheaper and better option than other approaches and strategies used to select from the large volume of web material often use a selective approach. The importance of policies to guide digital preservation is widely acknowledged and selection is included in the guidance given in the literature on creating a preservation policy.

## 2.4 Conclusions

The review of the literature has achieved a number of goals in order to fulfil objective 1. The review of the literature has revealed that the views of practitioners are underreported, as is their experience of selecting digital material. The theme of context has been apparent throughout the literature and it is possible to discern the differences between selection theory and practice in each domain for non-digital material. Archival appraisal is examined closely in the archive literature and there is a distinct body of theory deriving from practitioners working in the early and mid-twentieth century. With the advent of electronic records archives have had an impetus to examine theory and practice to take account of the need to appraise larger volumes of material. This has led to techniques such as macro or functional appraisal and documentation strategies. In comparison it can clearly be seen from the literature that the museum and library domains do not have the same theoretical basis for selection, instead basing practice on professional principles and knowledge of contextual factors, such as knowledge of users, the collections and the institution.

What is apparent is that whilst there are clear differences between domains for non-digital material, issues in digital material, such as copyright and resources, span domains. This chapter has examined a number of key practical management issues related to selection, including provision of adequate finance and resources, complex legal issues in particular copyright, and the need for policies to guide selection. Selection is commonly addressed through discussion of criteria (DPC Decision Tree 2006; Whyte and Wilson 2010 for example); these have been identified in the literature and compared. By synthesising the digital preservation literature it is possible to suggest a set of 'core' criteria which could be adapted for different contexts. These include cost, policies, use, value, IPR, documentation and technical ability. These are similar to those used for traditional selection, although the ability to preserve material, both technically and by having the appropriate resources and infrastructure, is a more prominent consideration.

Context as a theme is also apparent where the aims of different institutions within domains determine different selection activities. Within domains there may be differences due to scope and purpose; an academic library has different collecting priorities than a public library for example. There are differences between large and small institutions particularly in roles and responsibilities, where large organisations such as the BL and TNA have leadership roles, or the legal deposit role for the large deposit libraries. These types of institutions have greater capabilities to preserve digital material, due to increased resources and infrastructure, than smaller institutions. However there are advantages to co-operation in digital preservation, including sharing resources and infrastructure and this may be where their leadership and capabilities are particularly important.

The review of the literature has examined different terms that are used in each domain. Selection is used in libraries and describes part of the collection management function; it is part of a wider process of making appropriate material available to users. Much material which is preserved long term can be found in special collections, which may be treated as an archive. Accessioning happens in museums after acquisition and has particular implications of permanence; items are rarely de-accessioned. Items may be acquired for other purposes than preservation and these are not necessarily accessioned and so are viewed as being part of the main collection. Archival appraisal can happen before acquisition as well as after and is a process of determining value so material can be disposed of if no longer of value. In museums and archives in particular there are two processes; acquiring material and accepting responsibility for it permanently. Terms used in the digital preservation literature seem to be more often based on archival concepts, but 'selection and 'appraisal' are sometimes used interchangeably. This may reflect the cross-domain interest in digital preservation.

The theme of roles and responsibilities is also clear from the literature. Museum collecting has been more ad-hoc and controlled by individual curators to suit the needs of the museum. Traditional archival theory states that the archivist should not appraise records prior to their acceptance by the archive as this should be done by the creators, although more modern theory recognises the role of the archivists in creating the archive by appraising the material (Hedstrom (2002, p.37). Whilst ethically librarians should be non-biased in their selection (CILIP 2012b) and so create collections that reflect the needs of their users, museum curators have had more freedom to collect what they liked and so collections in museums can reflect individual interests. It is assumed that professional practitioners are still those best placed to select digital material.

This literature review has also revealed a number of assumptions which underpin selection. Concepts applicable to selection of digital material are few and based on assumptions that have not been widely examined. Some traditional assumptions are highlighted by Reed (2006, p.120) such as the notion of physical possession of an object, which is no longer a given in a networked environment, or that preservation can be dealt with separately from the creation of an object. The digital preservation literature assumes that obsolescence is an immediate threat to material (Waller and Sharpe 2006, p.8; DPC Handbook 2008, p.36; del Pozo, Stawowczyk Long and Pearson 2010, p.292 for example), though this assumption is becoming less prominent. In archives there is an assumption of permanent responsibility after the items have been accepted by the archive, similarly to museums, as these will be the most valuable items. But value is subjective and changes over time and between contexts and stakeholders; perspectives from the wider cultural heritage sector have helped to identify a wide variety of types of value that may be useful in selection. There is an assumption in the concept of significant properties, the specific characteristics of the material that must be preserved over time (Grace, Knight and Montague 2009, p.3), that value is discernible and similar for everyone, but this may not be so.

Selection, when considered in the digital preservation literature, is seen as mainly a practical issue, with the assumption that the theoretical basis is archival. There is focus in the digital preservation literature on the technical issues of digital preservation, with the OAIS model as common underpinning model. The concept of significant properties focuses too narrowly on the technical attributes of the material and also does not answer the question of 'significant to whom?' (See section 4.4.1 for the expert discussion of this concept which reflects this view). The concept of a 'designated community' is also used in the digital preservation literature, and is derived from the OAIS reference model, to determine who the material should be comprehensible to. This works well in defined areas such as institutional repositories but is less helpful when user communities are broader and less well-defined. Museums have a broad user base, from academics to children, and so might find this concept difficult to apply. Many types of libraries have a clear user community, such as academic libraries, although others serve a much broader community. It seems clear that digital preservation theory is based on archival theory, and it is assumed that as digital preservation is relevant to material held in different domains that these concepts are also applicable in all. Assumptions are made within the very small digital preservation community that are not necessarily shared by information producers and consumers who work in other domains (Gladney2008, p.7). This focus could distance digital preservation research from the communities that may be tasked with the responsibility for preserving material.

The following chapter will describe the research design and methods of sampling, data collection and data analysis for the different stages of this research.

## CHAPTER 3 RESEARCH METHODOLOGY

---

In chapter 1, the overall aim of the research was determined as: to investigate theory and practice of selection for digital preservation in UK memory institutions. This chapter describes how theory and practice will be investigated, the philosophy that underpins the research and the overall design including methods of collecting data. Questions of quality and ethics are then considered. The research has four elements - a literature review, preliminary expert interviews, practitioner interviews and policy analysis - and the specific methodology for each of these is described, including sampling, interview design and data analysis procedures.

### 3.1 Research philosophy

There are different concepts used in the literature when considering research philosophy. Kuhn argued that science is based on the accepted rules of the dominant paradigm which describe the 'legitimate problems and methods of a research field' (1970, p.10). A paradigm consists of '*some accepted examples of actual scientific practice..... [that] provide models from which spring particular coherent traditions of scientific research*' (1970, p.10). Researchers who share a paradigm share a particular set of rules and standards for practice. Lincoln and Guba (1985, p.33) characterise these rules as 'axioms'; these define the ontological, epistemological and methodological bases for different paradigms. Ontology refers to the nature of reality. Blaikie (1993, p.6) describes ontology as '*claims or assumptions that a particular approach to social enquiry makes about the nature of social reality.*' Even more broadly, Silverman (2006, p.13) states it is '*what reality is like and the basic elements it contains*'. Epistemology refers to the way we are able to know the world. Blaikie (1993, p.6-7) writes that epistemology '*refers to the claims or assumptions made about the ways in which it is possible to gain knowledge of this reality, whatever it is understood to be*'. Bryman argues that when considering an approach to research, epistemological issues are central to the framing of the research questions and the choice of methodology; they ask '*what is appropriate knowledge about the world?*' (2008, p.4). Silverman (2006, p.15) explains methodology as '*the choices we make about cases to study, methods of data gathering, forms of data analysis etc. in planning and executing a research study.*' Methods are the actual techniques used to collect and analyse data, such as interviews or questionnaires. Decisions on methods are considered within questions of methodology. Table 2 depicts the ontological, epistemological and methodological positions related to different research paradigms, based on explanations of research philosophies found in Bryman (2012), Robson (2011), and Lincoln and Guba (1985).

This research can best be described as having an anti-positivist stance, borrowing much from post-positivism. The nature of the aim and objectives means that the research is not focussing on the measurable nor is it seeking to verify hypotheses. Positivism cannot account for how people understand the actions of themselves and others, or how social reality is constructed because it focuses only on what can be observed. A positivist research philosophy was therefore rejected. This study takes account of individual understanding and experience; however if this research adopted a purely subjectivist approach, based solely on the perceptions of those doing selection, then many aspects of their context that act as influential factors would be missed or downplayed. This study considers reality to include not only observable phenomena but also abstract or hypothetical entities.

Table 2 Research paradigms

	Positivism	Post positivism	Interpretevism / Constructivism
Ontology (the nature of reality)	Reality is real and apprehensible	Reality is real and exists outside the mind	Multiple local and specific socially 'constructed' realities
Epistemology (how we can know reality)	Objectivist: findings are true; observable through our senses; aims for explanation and prediction through measurement	Findings are probably true, but reality is only imperfectly and probabilistically observable and apprehensible	Subjectivist: created findings; focus on understanding
Common methodologies (how we can investigate reality)	Experiments/surveys; verification of hypotheses, mostly quantitative methods	Case studies/interviewing; triangulation, interpretation of research issues by qualitative and some quantitative methods	Hermeneutical - researcher is a participant within the world being investigated

Having considered the assumptions of a number of perspectives, it is those associated with critical realism that underpin this research study. Critical realism is a meta-theory, rather than a method, based on the writings of Harré (1970) and Bhaskar (1975), developed further by other authors. Critical realism has an emancipatory and critical edge (Sayer 2000, p.18); it is critical because it does not accept at face value the accounts of social actors but criticises the practices and understandings which it studies. This research seeks to identify changes to assumptions and practice and so fits with this approach. A clear explanation of critical realism can be found in Danermark et al (2002) and Sayer (2000); a brief explanation of relevant aspects is presented here.

Critical realists argue that reality exists independent of our ideas of it and is objectively knowable. Reality consists of three different domains: the real, the actual and the empirical. The real is whatever exists whether we can observe it or not; the actual is what objects do whether we experience them or not; and the empirical is what we experience or observe them doing. The stratification of reality means there are more layers to a phenomenon than those that are observable. Real things have underlying structures, which Danermark et al (2002, p.47) define simply as '*a set of internally related objects*'. Mechanisms, which are based on these structures, act upon an object to produce an outcome; '*events arise from the workings of mechanisms*

*which derive from the structures of objects and they take place in geo-historical contexts'* (Sayer 2000, p.15). Many mechanisms can operate at once and affect each other (Danermark et al 2002, p.55-56). We need to know not only what actors do, but also about the structures and mechanisms that constrain and enable their actions, including the context in which they operate (Sayer 2000, p.25). In contrast to the positivist idea that experiences are the primary object of knowledge, in critical realism the appropriate objects of enquiry are mechanisms; the focus moves from the event to what produces the event (Danermark et al 2002, p.163). A further relevant aspect of critical realism is the conceptualisation of society as an 'open' system. We cannot manipulate social factors in order to study the effect of these manipulations; if we did this it would create a different social world that people would react to and the original would be lost. Regularities are only likely to occur in a closed, controlled system such as those artificially produced in an experiment. Social phenomena, such as heritage and selection, operate in an 'open' system which cannot be controlled in the same way as an empirical experiment. In an open system a causal mechanism can have many effects and different mechanisms can have the same effect.

To a positivist, reality is knowable through direct experience and observation, so truth would be the correspondence of observations to reality. In critical realism this is termed the 'epistemic fallacy' i.e. reducing reality to only our knowledge of it (Danermark et al 2002, p.205). For the critical realist there is a difference between the transitive dimensions of knowledge, which are the theories and discourses of science, and the intransitive i.e. the objects of science, the things that are studied (Sayer 2000, p.10). The intransitive world is not dependent on our transitive knowledge of it, which means that realist ontology can be combined with epistemological relativism, allowing for judgements of different explanations, theories and interpretations (Sayer 2000, p.47). The representation given of a phenomenon will always be from a particular position, which means there can be different descriptions and explanations of the same phenomenon. Each person's experience of selection and their understanding of the concept will be affected by their position in context and by unseen mechanisms and structures. In critical realism social phenomena are concept dependent and may constitute these phenomena as well as describe them (Danermark et al 2002, p.33). The example given is money - swapping metal discs is observable behaviour but we have to understand the concept and meaning of money to understand what is happening. Without the concept the behaviour would not happen. This means that the object of research is both real, in that it is made up of material practices, and also constructed, as social phenomena are concept dependent. Everyday concepts used by people can be revealing, whether they are 'right' or 'wrong'. Danermark et al (2002, p.200) clearly state that:

*We interpret the interpretations of other people. An understanding of significance and meanings is absolutely decisive for our ability to explain the social world....as social scientists we try to understand and explain what meaning actions and events have to people, but we also endeavour to produce concepts, which make it possible to transcend*

*common sense and attain a deeper understanding and explanation of a more abstract character.*

This, in summary, is the approach taken by this research. This research will seek to explore underlying mechanisms and structures that affect selection through the analysis of the subjective accounts of social actors and consideration of their contexts.

## **3.2 Research design**

Research designs are frameworks for the collection and analysis of data (Bryman 2004, p.26). A particular design is also related to the type of criteria that may be used to evaluate the research. This section will consider how data will be collected and analysed, how quality may be considered, and the ethical issues involved.

### **3.2.1 Qualitative research**

There are two broad categories of research methods: qualitative and quantitative. Danermark et al (2002, p.163) argue that the traditional division between quantitative and qualitative research is 'limiting and misleading'. However many research methods textbooks contain a description or comparison of quantitative and qualitative research (see Silverman 2006, p.35 for example). As Bryman points out, the distinction is still useful as it provides shorthand terms for a range of issues; it also helps classify social research methods (2004, p.19).

Creswell (2009, p.17) gives a useful comparison between characteristics of quantitative, qualitative and 'mixed methods' strategies. 'Mixed methods' is usually characterised as a combination of qualitative and quantitative strategies. Quantitative methods use the positivist (or post-positivist) experimental approach to discover explanations, test hypotheses and gather facts about the world. Qualitative research attempts to see the world from the point of view of participants in it and to contribute to the development of knowledge about it. It encompasses a variety of paradigms, such as post-positivism, constructivism and critical theory and an even wider variety of methodologies. It is difficult to generalise about such a broad range, but Miles and Huberman (1994, p.6-7) and Robson (2011, p.19) list recurring features of qualitative research, so by combining this with the comparisons Denzin and Lincoln (2005, p12) make with quantitative research we can see that qualitative research has the following features:

- conducted through intense or prolonged contact with a 'field', a natural setting, which is often reflective of everyday life
- the researcher's role is to gain an integrated or holistic overview of the context under study
- the researcher attempts to capture data on the perceptions of local actors
- the researcher is concerned with securing rich descriptions of the social world
- a main task is to explicate the ways people understand or account for their actions and situations – there is a focus on meanings

- the researcher is the main ‘instrument’ – objectivity is not highly valued
- most analysis is done with words, such as historical narratives, first person accounts, or biographical and autobiographical material (Denzin and Lincoln (2005, p12).

Within critical realism the division between qualitative and quantitative is termed intensive and extensive. In table 3 the comparisons between intensive and extensive research are used to define each type (table modified from Danermark et al 2002, p.165):

Table 3 Intensive and extensive research

	Intensive	Extensive
Typical research questions	How does a process work in a small number of cases? What produces a certain change? What did the agents actually do?	What are the regularities, patterns, or distinguishing features of a population? How widely are certain characteristics or processes distributed?
Typical methods	Study individuals in their causal contexts; interactive interviews; ethnography; qualitative analysis	Large scale survey of population or representative sample; questionnaires; standardised interviews; statistical analysis
Limitations	Unlikely to be representative, average or generalisable.	Limited explanatory power.

Whilst the authors are careful to say that ‘the research process involves an intensive and an extensive element’ (Danermark et al 2002, p.167) it is clear that they are much more in favour of an intensive approach. Sayer (2000, p.21) describes critical realism research as contextual, intensive (looking at a small number of cases in depth) and anti-reductionist (looking at many aspects or dimensions of a problem). This research utilises an intensive approach. In addition to dealing with what (descriptive) and why (explanatory) questions, this research seeks an explanation of the ‘how’ question. It aims to uncover underlying mechanisms and structures that explain unobservable phenomena through the analysis of the subjective accounts of social actors and supporting organisational documentation. When the context and individuals’ accounts are important, then qualitative (intensive) methods are more appropriate for the richness of data required. Therefore the decision was made to address the objectives using a qualitative approach, as this is more appropriate to the research objectives and philosophy.

Within the broad category of qualitative methods, there are many further approaches that could be employed in research. Table4 (adapted from Thomas 2006, p. 241) below summarises four often used qualitative approaches to data collection and analysis, focusing on the types of questions posed, methods of analysis, and the outcomes that each may produce and how those findings are commonly presented.

Table 4 Qualitative methodologies

	General Inductive Approach	Grounded Theory	Discourse Analysis	Phenomenology
Analytic strategies and questions	What are the core meanings evident in the text, relevant to evaluation or research objectives?	To generate or discover theory using open and axial coding and theoretical sampling	What meanings are created through the social practices of language? Language constitutes and reproduces social worlds (Bryman 2012, p.528)	Seeks to uncover the meaning that lives within experience and to convey felt understanding in words
Methods of analysis	Development of codes then categories from the raw data most relevant to research objectives	Initial coding Create a series of categories Determine the core category which integrates the data	Close examination of texts and the structure of language use Interested in how the story is told	Identify descriptions of the phenomenon. Coding holistic, selective, and/or detailed. Describe the essence of the phenomenon.
Outcome of analysis	Themes or categories most relevant to research objectives identified	A theory that includes themes or categories	Multiple meanings of language and text identified and described	A description of lived experiences
Presentation of findings	Description of most important themes	Description of theory that includes core themes	Descriptive account of multiple meanings in text	A coherent story or narrative about the experience

It is clear that using discourse analysis in this research is incompatible with the research objectives and the underlying research philosophy. For realists interpretation of meaning involves not only the actors' discourse but also the context of that discourse, and

'unacknowledged conditions and things that can happen to people regardless of their understandings' (Sayer 2000, p.20). However a phenomenological approach was considered for this research. Creswell defines phenomenology as '*a strategy of inquiry in which the researcher identifies the essence of human experiences about a phenomenon as described by participants*' (2009, p.13). Phenomenology focuses on sense experiences as the source of knowledge. The focus on experience rather than the underlying structures and mechanisms differs from the critical realist approach favoured by the researcher. Additionally, the objectives of this research are to understand aspects of selection but not to identify its 'essence'. A more explanatory and less descriptive approach is more appropriate to fulfil the objectives.

Consideration of the aim of the research also excluded a purely grounded theory approach; the aim of the research is not to generate theory but investigate and explore. Grounded theory as a research strategy was first developed by Glaser and Strauss, in their book '*The Discovery of Grounded Theory*' (1967), written as a reaction to the dominance of positivistic quantitative methodologies used in contemporary sociology (1967, p.17). Their original version of grounded theory proposed systematic strategies for qualitative research which included the following defining components, as summarised by Charmaz (2006, p.5-6):

- Data collection and analysis to be performed simultaneously
- Codes and categories are formed from considering the data and are not imposed from preconceived hypotheses
- The use of the constant comparison method
- Theory develops as the process continues
- Memo writing as an integral part of the analytic process
- Sampling dictated by theory construction
- Conducting the literature review after the analysis

Grounded theory has been further developed in different directions, reflecting divergent views of the two original authors, though it is clear that the approach has been very influential in qualitative research; Charmaz (2006, p.6) describes it as legitimising qualitative research as a credible approach to research. However Layder (1993, p.61) argues that grounded theory is limited by its insistence on focusing on only what can be observed - it is limited by the data rather than guided by it. This limits understanding as it cannot tell us about the underlying unobservable mechanisms that produce the behaviour, nor takes account of the wider context.

Layder (1993, p.62) goes on to argue for a form of grounded theory that incorporates a greater variety of strategies and theoretical perspectives and includes wider contextual information, thus incorporating ideas from grounded theory in a realist approach. It is clear that, as Robson (2011, p.150) states, there is no '*basic incompatibility*' between grounded theory and realism as it offers helpful guidelines. Robson goes on to say that '*if the guidelines are used as flexible*

*tools rather than rigid rules, grounded theory gives researchers a broad method with distinct procedures that work in practice'* (2011, p.150). However the aim of this research is not solely to build theory but to explain and understand selection, and also considering the resource constraints, it was felt pragmatic to use a flexible approach, based partly on the 'general inductive approach' described by Thomas (2006) outlined in table 4, and in particular on the guidance from Miles and Huberman (1994), who make clear that they consider themselves 'realists' (1994, p.4) but have a practical approach to analysis. For them, qualitative research and analysis is in three concurrent activities: data reduction, data display and conclusion drawing. These will be discussed further in consideration of the methodology for each phase of this research in the rest of this chapter. This researcher agrees with Robson (2011, p.38), who classifies himself (somewhat facetiously) as '*realism-lite*', in that he agrees with many aspects and ideas of critical realism, but allows himself to go forward '*pragmatically selecting ideas and terminology from different realist approaches which appear likely to be useful...*'.

### **3.2.2 Data collection**

This section will give an overview of the methods used to collect data in this research. The specific processes and the methods of analysis are detailed later in the sections on each phase of the research.

Miles and Huberman (1994, p.6-7) present two pictorial taxonomies of research methods, which illustrate the dozens of methods available to the qualitative researcher, categorised by method or the purpose of the research. The methods chosen from the large number of possibilities reflect the ontological, epistemological and methodological axioms already discussed and aim to fulfil the research objectives.

In reflecting on the aims and objectives of this research, a questionnaire to determine the extent of digital preservation within memory institutions was considered. This method was ultimately dismissed for a number of reasons:

- Potential problems with participation levels. (Bryman2004, p.135) discusses this problem)
- Difficulty in asking further questions or collecting additional data which provide the in depth data necessary to fulfil the research objectives
- The literature review had also uncovered much other work investigating aspects of digital preservation activities so it was felt that yet another questionnaire survey was not the best use of the researcher's time
- The choice of a qualitative approach precludes the use of questionnaires which provide quantitative data. Questionnaires with open questions inviting written answers would be

more appropriate, though they suffer from problems also, such as a reluctance of respondents to write long answers (Bryman 2004, p.134).

Other methods of data collection commonly associated with qualitative research include focus groups and observation. Focus groups may be defined loosely as interviews carried out in a group setting (Robson 2011, p.293). More specifically, Bryman (2012 p.503) describes them as a form of group interview with several participants, focused on a topic or a number of topics, which is partly interested in the interaction of the group as well as the content of their answers. This method of data collection would have been suitable for this research; however problems would have arisen with arranging a suitable time and venue for a group of people to meet. This is not possible especially for the first phase of expert interviews where the aim was to gather data from geographically dispersed individuals, from different cultural heritage domains. Whilst it may have been possible to gather the experts together at a conference or other such event, there was little likelihood of this happening within the timescale that was available for this work. The researcher kept in mind the possibility of using a focus group to gather data from the practitioners, but again geographic distribution and timescales have not permitted this option.

The interviewer recognises that the Delphi method may have been an appropriate method of conducting the expert interviews in particular. This has been defined by Linstone and Turoff (1975, p.3) as: '*...a method for structuring a group communication process so that the process is effective in allowing a group of individuals as a whole to deal with a complex problem*'. General features of this technique are that a panel of experts are convened (though they do not need to meet) and sent questionnaires. The results are analysed and are distributed back to the panel. This process continues until a consensus is reached, though a modified Delphi study may allow for a divergence of opinion. The Delphi method is often, though not exclusively, used for prediction purposes. Whilst this technique has the advantage that it can be conducted remotely, it may take a long time as many iterations of questions may be needed. It was therefore felt that although these techniques had promise, the time constraints on the research, especially for the initial expert interviews, were too great to allow this to be used. The likelihood of the digital preservation experts being willing to take part in a longer study were considered low.

Observation is a commonly used technique in an ethnographic approach, although it is also appropriate to other approaches, such as phenomenological studies (Langdrige 2006, p.80). Silverman (2006, p.67) explains that observation is a method of gathering data that involves looking, listening and recording. In this research, the researcher does not consider that observation of selection taking place would reveal any useful data, considering the stated aim and objectives.

After careful consideration, it was decided that this research would be conducted through a comprehensive literature review, in-depth interviews with stakeholders and an analysis of policy

documents from cultural heritage institutions. For the policy analysis phase, data were gathered through requesting policies on appropriate mailing lists and through searching the internet. This will be described further in section 3.6.

### 3.2.3 Interviews

Interviews are a particular form of professional conversation (Kvale 1996, p.5). Conducting successful interviews is a craft, a skill to be learnt, and there are few rules as to how to go about them. Interviews are an appropriate method of data collection in this research as the objectives refer to opinions, perceptions and assumptions as well as facts, which can easily be gathered through interviews with respondents. Table 5 below summarises advantages and disadvantages of interviews, as described by Silverman (2006, p.118) and Robson (2011, p.279). These are general observations and may not apply to all interviews as they can vary according to structure and method of delivery. The type of knowledge that is gained from an interview is very much related to the axioms underlying the research and the theoretical viewpoint of the researcher. Positivists would use interviews as a way of discovering ‘facts’ about the world; constructionists would ask how meaning is constructed through an interview (Silverman 2006, p.118). In this research, the first phase of interviews used the experts as a ‘resource’, and aimed to elicit their views on what are important issues in selection for digital preservation. The second phase with practitioners also used them as a resource to find out how selection is done in their institution. But in line with the epistemology and the ontology described earlier, these interviews are also seen as a way of identifying the underlying mechanisms and structures inherent in the context in which the practitioners operate, as interpreted, experience and believed by them.

Robson (2011, p.279-280) indicates that there are three broad types of interviews, differentiated by their structure or standardisation:

- Fully structured interview – characterised by predetermined questions and fixed wording delivered in a pre-set order
- Semi-structured interview – the interviewer has an interview guide of topics or questions to be covered but can modify these according to the flow of the interview
- Unstructured (open) interview – there is a lack of predetermined questions; the interviewer has general areas of interest which serve as starting points for the conversation

In this research open interviews were conducted with the experts, and semi-structured interviews with the practitioners. The details of these can be found later in the description of each phase.

Table 5 Interviews

Advantages	Disadvantages
The interview format is flexible and adaptable	The potential interviewee must be able to find time to take part
People are relatively familiar and comfortable with the format	Can be time-consuming to perform and transcribe; May produce a very large quantity of data
People like talking about their work and lives, thus giving a richness of data	Potential problems of bias
Able to understand non-verbal communication	Travel to where the interviewees are based may be expensive and time consuming
Able to follow up interesting responses at the time	Potential problems with the equipment, batteries and background noise, especially if the interviewee prefers to meet in a public place such as a cafe
	Transcriptions of interviews can erase context (but notes taken during and immediately after the interview help combat this)

In addition to the structural variations found in interviews they may also vary by the method of delivery. The aim throughout this research was to conduct interviews face to face with the respondents, to take full advantage of the non-verbal communication this affords. However this has not always been possible so some interviews were conducted over the telephone and some via email, due to geographical considerations and time constraints by busy participants. Whilst telephone interviews are straightforward, despite the lack of face to face non-verbal communication, email interviews are more complex. They can be conducted in a variety of ways, including asking questions one by one in a 'conversation' format where the interviewer waits for an email response before asking another. Another possibility is to email all the questions at once, though this has the potential to become more like a questionnaire, unless the respondent allows the interviewers to ask follow up questions and is prepared to give full answers. The latter format was used in this research and the respondents were often very generous with their responses. The researcher obtained agreement to ask further questions if necessary. Robson (2011, p.290) describes advantages and disadvantages of both these methods of delivery, some of which are highlighted in table 6.

Table 6 Telephone and email interviews

Telephone interviews		Email interviews	
Advantages	Disadvantages	Advantages	Disadvantages
Quick and cheap	Often shorter than the equivalent face to face interview	Cheap and quick (no transcription necessary)	Can take a long time to respond
Potential reduction in bias due to interviewer responses	Lack of visual cues	Able to do more than one interview at a time	Lack of non-verbal communication
	Lack of contextual information	Can reflect on follow up questions and the respondents are able to reflect on their responses	Impersonal - there is little opportunity to build rapport

### 3.2.4 Quality in qualitative research

Positivist research has clear criteria of reliability, validity and objectivity. These can be defined as follows:

Table 7 Positivist quality criteria

Reliability	indicates the ability of a test or measurement to be replicated or yield consistent results (Silverman 2006, p.282)
Validity	the ability of the research instrument to measure what it is intended to measure (Gray 2009, p.155)
Objectivity	When the inquiry is 'value-free' - the researcher does not allow their own values impact on the research; the researcher maintains distance from the object of research (Lincoln and Guba 1985, p.300)

According to Robson (2011, p.155), these measures have become operationalised within quantitative research and the challenge is to adapt them to the conditions of qualitative designs. The concepts of reliability and validity can be problematical in qualitative research as they are based in the positivistic paradigm (Silverman 2006, p.302) which this research has rejected. The social world is an open system and as such is not controllable in the same way as an experiment in the natural sciences. For example, as Robson (2011, p.155) explains, it is not possible to exactly replicate the circumstances of a social event or setting in order to replicate findings. There is no agreement in the research literature on the appropriate criteria to use; the debate has been fierce as it goes to the heart of the question of credibility of qualitative research as a strategy (Robson 2011, p.155; Miles and Huberman 1994, p.277).

Despite these issues, it is important for qualitative researchers to consider both reliability and validity (or their equivalents) in their research designs. Multiple versions of lists of criteria of

measures for assessing qualitative rigour are put forward in the research methodology literature (such as Gray 2009, p.194; Bryman 2008, p.377), and these criteria are linked to the paradigm to which a researcher adheres. As Lincoln and Guba (1985, p.294) state '*different basic beliefs lead to different knowledge claims and different criteria*'. A further difficulty is that multiple words meaning the same type of measure are used by different authors. The suggested criteria from Miles and Huberman (1994, p.278-279) includes elements to ensure the rigour and quality of the research, whilst combining similar terms. These strategies have been successfully combined with those suggested by Robson (2011) and Silverman (2006) to ensure quality and rigour. These are shown in table 8 on p.74.

### **3.2.5 Ethical considerations**

As the research involves human participants, ethical issues must be considered. A Loughborough University ethical clearance checklist was completed prior to interviews beginning. This did not identify any areas of concern and no further ethical clearance was required, though it prompted careful consideration of compliance with Loughborough University Ethical Advisory Committee's Code of Practice on Investigations Involving Human Participants. Considerations include:

- The interviewees are all adults and are not in vulnerable groups.
- Informed consent for the interviews was acquired from the interviewees and they were informed they could withdraw from the interview at any time.
- No financial incentive was given to take part.
- The purpose of the research and details of what the interviewees could expect if they agreed to participate was made clear.
- Locations for the interviews were selected in full consultation with the interviewees.
- Anonymity is ensured by making every effort to keep their responses anonymous within the thesis; the experts and practitioners are only identified by a number and care is taken with quotations, so no identifying detail is present.

Table 8 Techniques to ensure quality

Criterion	Definition	Methods employed in this research
Objectivity/ confirmability	Freedom from bias or explicitness about bias	Being explicit in the theoretical assumptions underlying the research  Describing the methodology in detail so the process of the research can be followed and it is transparent  Keeping a 'research diary' and using memos, so an 'audit trail' is created of thoughts and activities
Reliability/ dependability/ auditability	The degree to which the research process is consistent, clear and stable across time and methods	Making clear the philosophical stance taken by the researcher  Recording respondents words verbatim  Transcribing carefully to provide an accurate rendering of respondents words  Keeping records (as above) of activities  Checking, and writing memos about, the codes and their meanings
Internal validity/ credibility/ authenticity	Findings should make sense and be credible	Presenting quotes in the report from respondents  Use of constant comparison methods during analysis - comparing data to data and case to case  Using appropriate tabulations (Silverman 2006, p.299; Miles and Huberman 1994, p.252 - qualitative research does not mean excluding simple counting techniques to indicate variance or prevalence
External validity/ transferability/ fittingness	The extent to which the results have a larger import, are transferable to other contexts and whether they 'fit'	Provide comprehensive information on the context in which the research is carried out; provide a 'rich' description  Careful and thoughtful sampling; being explicit as to method of sampling  Being explicit about areas of uncertainty  Presenting work for peer review through writing for publication
Utilisation/applic ation/action orientation	The usefulness of the research and who may benefit from it	Ethical concerns are clearly addressed  Suggestions for further research

### 3.3 Literature review

The resources for this review were drawn from the library and digital preservation literature along with wider cultural heritage material, including relevant archive and museum resources. By searching both general internet resources and specialist databases an extensive range of literature was found and analysed. Initial keywords for searches were identified and then potential follow on keywords were identified throughout the literature review process. Derivatives of all these keywords were used for searching and are shown in table 9 below. The core keywords were searched for either singularly or in combination using Boolean commands, such as AND and OR, and truncations.

A variety of tools were used to find the literature, such as library catalogues, bibliographies, abstract and indexing services, Internet search engines and discussion lists. These included Library and Information Science Abstracts, Emerald EMX95, the Theses Index, the digital preservation discussion list on JISCMail (JISC n.d.b) and Google Scholar. The Google search engine produced too many results, even with the use of Boolean search terms, so Google Scholar was used to help narrow the number of returned hits and focus on scholarly works. World of Science and Zetoc were searched early on but were not found as useful. A number of useful websites were also identified, including the Department for Culture Media and Sport (DCMS), Collections Link, the Digital Curation Centre (DCC), the Digital Preservation Coalition (DPC), Joint Information Systems Committee (JISC), Preserving Access to Digital Information (PADI) and the Council for Library and Information Resources (CLIR). These sites had many links to other relevant material. The literature search was run more than once and automatic email updates were subscribed to for the most useful journals to keep up to date with new articles about selection and digital preservation.

Table 9 Search terms

Core keywords	Derivatives	Follow on keywords
culture	cultural; cultur*	
heritage		memory
selection	select; select*	appraisal
library	librar*; libraries	public
museum	museum*	social
preservation	preserve; preserv*	capital
value		collection
digital	digitisation; digit*	ethics
archive	archiv*	technology
curation	curat*	acquisition

### **3.4 Expert interviews**

#### **3.4.1 Sampling**

This phase of the research employed purposive sampling which aimed to recruit recognised experts in the field. Purposive sampling is the process of the researcher choosing respondents deliberately in relation to the research questions or theoretical considerations (Robson 2011, p.275). The experts were selected on the basis of reputation and body of work in the field of digital preservation, as found during the literature review. This research does not depend on a statistical sample that attempts to be representative of any population. It does however require qualified experts who have a wide understanding of the issues. As Robson describes, searching for mechanisms initially may mean ‘...you explore with those who are knowledgeable about the setting which mechanisms and contexts appear to be the ‘bankers’” (2011, p.35).

Initially ten experts were contacted. Through this initial contact and recommendations from some of those contacted, eventually eight experts agreed to share their views. Six of them were from the UK, one from Canada and one from Australia. One interview was conducted via email, two via the telephone and five face to face. The telephone and face to face interviews were recorded digitally. The use of a digital recorder meant the researcher was able to concentrate better on what the interviewee was saying, and respond and probe deeper where appropriate, rather than just writing down the words.

The experts who took part in this phase of the research included academics, consultants and practitioners who are considered experts or leaders in the field; many have published or spoken extensively on the topic of digital preservation.

#### **3.4.2 Interview design**

The first phase of interviews consisted of open interviews, which were not completely unstructured. There are a number of initial (sensitising) topics the researcher was interested in particular, such as value and roles, which a completely unstructured interview may not have addressed. This approach gave interviewees flexibility in how they answered and allowed the interviewer to probe areas and follow up issues as necessary.

Following the literature review a number of sensitising concepts were identified:

- Value
- Roles and responsibility
- Difference – between institutions and between types of material

The term ‘sensitising concept’ was first described by Blumer (1954). These are not definitive concepts with precisely defined meanings, but are concepts that ‘gives a user a general sense

of reference and guidance' and 'merely suggest directions in which to look' (Blumer 1954, p.7). They can be improved and refined as research proceeds. As Blaikie (2000, p.137) states '*While a researcher needs some guidance, it is necessary to view the research itself as a process in which meanings of concepts are developed.*'

These sensitising concepts formed the basis of the interview questions. A short set of open interview questions, with suggestions for follow up questions, were devised:

- What do you think are the main similarities or differences between preserving traditional and digital material?
- What do you think are the main similarities or differences between the approaches to selection (of any type of material) in libraries, archives and museums?
- What do you think are important influences on selection decisions?
- What does the concept of value mean in the preservation context? How can it be determined?
- Who is/could be/should be responsible for selection and for determining the grounds for selection? Whose role is it? Is it an ethical issue?
- Who do you think are the main stakeholders (interested parties) in selection?
- Are there any other issues that you think are relevant to selection for (digital) preservation?

The face to face interviews were held in locations chosen by the respondents, and included a pub, their offices and a cafe. Due to the open nature of the interviews, the interviewee was able to explore the issues contained in the questions in their own way. The number of questions was kept intentionally few to allow the interviewee to talk in depth about their opinions and experiences and for the researcher to probe and question further where clarification was required. The interviews were constrained by the busy schedules of the experts; one interview lasted only half an hour with others lasting between three quarters of an hour and an hour and three quarters.

### **3.4.3 Data analysis**

The interviews were transcribed manually, in full, though without including the features of speech such as how words are pronounced, pauses, overlaps and so on, as this level of detail is not necessary for the approach taken in this research. The transcripts were then coded using the Atlas.ti software. This software was chosen as it is supported by Loughborough University; it is widely used and there is helpful literature in how to use it easily available; and the licence costs are not prohibitive. The outcome of this stage is a descriptive analysis of the issues

identified as important in selection by the experts, with various sections related to the categories (or pattern codes) created.

A 'general inductive approach' is described by Thomas (2006) and has much in common with the approach to qualitative analysis described by Miles and Huberman (1994) and in grounded theory (Charmaz 2006; Strauss & Corbin, 1998). Thomas (2006, p.238) explains that the purposes underlying the GIA are:

- Condense large and varied amounts of raw data
- Establish clear links between objectives and findings derived from the data
- Develop a model or theory about the underlying structure of experiences or processes that are in the raw data

This reflects the three activities Miles and Huberman employ (1994, p.10-11):

- data reduction - the process of 'selecting, focusing, simplifying, abstracting and transforming' the raw data;
- data display - 'an organized, compressed assembly of information that permits conclusion drawing';
- conclusion drawing and verification.

The process of inductive coding is described by Thomas (2006, p.241):

- Preparation of data - 'cleaning' the data so that files are in common formats, printed, backed up and so on.
- Close reading of the text to gain familiarity
- Creation of categories (codes) by labelling meaningful segments of text. Text can be left uncoded or coded more than once.
- Continuing refinement and revision of the emerging coding system

Miles and Huberman (1994, p.58) recommend beginning coding with a provisional set of codes. However at this stage in the research a more inductive approach was used, which suited the exploratory nature of the expert interviews. Initial coding is created by defining what is seen in the data and codes emerge through close reading of the data. It is a very iterative process of comparing data with data (Charmaz 2006, p.46). Charmaz recommends coding word by word or line by line (2006, p.50), though in this research the text segments were chosen rather by their meaningful content than any predetermined idea of size. The sections of data that were coded varied in size, anything from a phrase to a paragraph. Definitions of codes were developed as the coding process continued and memos were written about codes, reflections and observations about the data by the researcher. Memoing is a very useful method of recording

ideas and potential concepts that occur to the researcher as they are going along and continues all through the analysis.

The result from the initial coding may be a long list of descriptive codes summarizing the data but the next level in the coding process is what Miles and Huberman term 'pattern coding' (1994, p.69). This relates to the fourth stage as described by Thomas above, of refinement and revision. It is a way of grouping initial, descriptive codes into 'explanatory or inferential codes, ones that identify an emergent theme, configuration, or explanation' (Miles and Huberman 1994, p.69). Pattern codes are generated by 'looking for threads that tie together bits of data'. Descriptive codes have little intellectual shape on their own, but after coding the interviews and initial memoing, re-reading of codes, memos and comments was conducted and some more common themes began to emerge. By revisiting the initial coding the codes started to take on more interpretive elements, rather than purely descriptive. Codes were displayed in the 'network view' available in Atlas.ti, and also printed out onto paper, which enabled a visual representation and organisation to be used to cluster codes into 'families', or categories, forming meaningful groups in order to create a more conceptual coding scheme (Miles and Huberman 1994, p.63).

### **3.5 Practitioner interviews**

Following on from the interviews with experts and concurrently with the policy analysis phase a series of interviews with practitioners in libraries, archives and museums were undertaken. The aim of these interviews was to gather the views of people who are directly concerned with digital material, or the issues surrounding digital material, in their day to day work. This section will describe in detail the methods of data collection used in this phase, along with the methods of analysis and results to date.

#### **3.5.1 Sampling**

The types of sampling used in this phase of the research may be described as a mix of purposive, snowball and convenience sampling. These are all types of non-probability sampling which are appropriate to qualitative research. Purposive sampling has already been described in section 3.4.1. Snowball sampling can be regarded as a type of purposive sampling where the researcher identifies a few individuals in the population of interest, and then these individuals identify or recommend further individuals that could be included in the study. Convenience sampling is choosing the nearest, most convenient respondents (Robson 2011, p.275-276). Here the initial sample has been identified through a preliminary analysis of potential stakeholders to identify those who may have an interest in the topic. Potential participants were found through:

- MLA designated collections list
- Society of Archivists
- Culture24 listings of institutions
- JISC digital media website
- Other interviewees
- Google
- Respondents to request for policy documents
- Met at a conference.

It was recognised that the qualitative approach taken in this research would generate large amounts of data. This needed to be taken into account in sampling and it became appropriate to limit the range of stakeholders to be included in the research in order to adhere to the scope and time limits of the research project. Thus the sampling focussed on the people with direct day-to-day involvement with the issues raised in the previous interviews, i.e. curators, librarians, archivists and their managers. The differences between managers and practitioners was not always clear; there are different levels of management, such as high level strategic managers and lower level line managers still with direct contact with day to day practitioner work. This difference was not focussed upon during the sampling for this phase.

Email templates were devised to request interviews and these were used and adapted to each individual potential participant. The actual numbers of interviews requested and performed are shown in table 10.

Table 10 Practitioner interview numbers

	Archives	Museums & Galleries	Libraries	Total
Requested	18	13	15	46
Accepted	11	5	9	25

The breakdown of the type of respondents identified as stakeholders in the preservation process which were actually interviewed is shown in table 11.

Table 11 Practitioner interviews by setting and role

Completed	Archives	Museums & Galleries	Libraries	total
Managers	4	3	4	11
Non-manager practitioners	7	2	5	14

Table 12 on p.82 describes the interviewees in more detail; A – Archive; M – Museum; L – Library.

### 3.5.2 Interview design

The intention was to formulate interview questions separately for each group, however as sampling continued and it became clear that the interviews would concentrate on practitioners and managers the decision was made to formulate one set of questions closely based on all of the identified issues. This could then be adapted on a case by case basis and some supplementary questions targeted at either of these roles were created, to be used when appropriate within an interview and with consideration of the interviewees' responses.

The interviews were semi-structured as this approach allows questions to be asked on particular topics but also gives interviewees flexibility in how they answer and allows the interviewer to probe areas and follow up issues as necessary. A copy of the interview schedule is shown in table 20 in appendix 4. This allows for consistency across the cases, ensuring appropriate information is collected regarding the research questions. However they are flexible enough to allow the interviewer to probe any areas of particular interest, or to allow the interviewee to take the discussion in a direction most relevant to them. All the interviews began with an introduction by the researcher to themselves and the purpose of the interviews, allowing the researcher to make sure the respondents were still happy to continue with the interview. The shortest interview lasted thirty five minutes and the longest an hour and forty minutes, with the majority lasting about an hour and twenty minutes.

Table 12 Practitioner interviewees

code number used in text	domain		description
1	A	manager	digital preservation manager at a large public archive
2	M	practitioner	curator at a national museum
3	L	manager	manager at a national library
4	A	practitioner	archivist at a company
5	M	practitioner	curator at a national museum
6	L	practitioner	librarian at a data repository
7	A	practitioner	archivist at a local authority
8	L	manager	manager at a national library
9	A	practitioner	officer at a national archive
10	A	practitioner	archivist at a large archive
11	A	practitioner	curator at a data service
12	A	practitioner	archivist at a local authority archive
13	L	manager	digital manager at a large library
14	A	manager	local authority archivist
15	A	manager	manager at a national archive
16	L	practitioner	librarian at a special library
17	L	practitioner	academic librarian
18	L	practitioner	librarian at a local authority library
19	L	manager	manager at a large university library
20	A	manager	university special collections archivist
21	M	manager	manager at a national museum
22	M	manager	manager at a national museum
23	M	manager	curator at a local museum
24	L	practitioner	information officer at a local authority
25	A	practitioner	archivist at a large company

### 3.5.3 Data analysis

As recommended by Miles and Huberman (1994, p.58) an initial broad coding framework was used to analyse the practitioner interviews. This was based on the themes and issues identified in the expert interviews; the questions asked of the practitioners were based on these themes and so the responses clearly relate to these areas of concern. The coding framework included:

Table 13 Initial practitioner coding framework

Professional issues	Institutional issues	Technological issues	Conceptual issues
Roles and responsibilities	Roles and responsibilities	Loss	Significant properties
Ethics	Risks	Format	Value
Knowledge and skills	Legal issues		The future
Current users	Policies		
	Finance and costs		

However the researcher has been careful to be open to other potential themes and to creating other codes and categories as issues are found in the data, hence the analysis proceeded inductively. Transcripts were coded both top-down using the previously identified coding scheme, and also bottom-up with each theme fleshed out with more detail and new codes added. Each interview has been coded and the codes brought together into categories. Findings from the practitioner interviews can be found in chapter 5.

### 3.6 Policy analysis

This research examined the digital preservation policy documents from a number of libraries, archives, museums and other organisations concerned with the preservation of digital cultural heritage material. A factor to arise from the expert interview phase is policy, as described in section 4.2.4. It therefore seemed practical to investigate this aspect further not only through questions in interviews with practitioners but also through the documentary sources. These include the literature on digital preservation policies and the policies themselves. This work was conducted concurrently with the practitioner interview phase. The documentary policy analysis will be presented in chapter 5, section 5.5.5.

### 3.6.1 Sampling

Purposive sampling is the process of the researcher choosing respondents deliberately in relation to the research questions or theoretical considerations (Robson 2011, p.275) and a form of purposive sampling was used, along with convenience sampling. The latter was appropriate because the relevant policies may be internal organisational documents and not available publically.

In order to gain access to these documents, the researcher posted a request for policies on a number of JISC mailing lists. The text of the request was designed to be quite general and friendly, and included information about the researcher and what the purpose of the request was. Links to eleven policies from people working in the institutions below were received:

- Archaeology Data Service (Austin and Richards 2009)
- Northumberland Collections Service
- Wellcome Library (Checkley-Scott and Thompson 2007)
- Hampshire Archives
- National Library of Wales
- Parliamentary Archives (Brown 2009)
- UK Data Archive
- DetKongeligeBibliotek - The Royal Library (Denmark)
- University of Edinburgh Library
- DSpace@Cambridge (The University of Cambridge institutional repository)
- State and University Library, Denmark

One further response from the National Library of New Zealand was of their draft policy which was not to be made public. In addition policies from the following institutions were sent to the researcher via email:

- Greater Manchester Archives
- The Royal Commission on the Ancient and Historical Monuments of Scotland
- The Earth Resources Observation and Science (EROS) Center (USA )
- West Yorkshire Archives Service

In total sixteen policies were obtained in this way. Many further responses were received from people kindly providing references to various online resources concerning preservation policy and policy formulation.

In addition, further publically available policies and guidance on policies were searched for on the internet and sampled purposively. This produced six more policies from the following institutions:

- London Metropolitan Archives

- State of Victoria Library, Australia
- National Library of Australia
- Libraries and Archives Canada
- National Museum of Australia
- Yale University Library

There was a conscious effort to include policies from outside the UK to provide a broad sample of policies. The main difficulty was selecting policies that related to preservation, as institutions may have many policies relating to the same area, for example a collecting policy, an acquisitions policy, a preservation policy; the names for these policies are used interchangeably and the information specifically about digital preservation can be hidden within other policies. Twenty two policies have been included in the sample in total. As can be seen from the lists above the policies included in the sample are from reputable organisations.

### **3.6.2 Data analysis**

The documentary analysis includes an examination of the policies sampled and a comparison between the policies and recommendations from policy guides found in the literature to determine the clauses and concepts suggested in the literature and those used in practice. This is a qualitative analysis of the content of the policies and policy guidance documents; a quantitative content analysis would not be appropriate for the exploratory nature of the research, the intensive approach taken and be outside the scope of the research.

Firstly there was an analysis of recommendations found in policy guides for clauses, or provisions, to be included in digital preservation policies; this is included in the review of the literature in section 2.3.6. The policy guide literature was analysed by means of comparison and counting to produce an overall taxonomy of clauses. The policy guides were read carefully, taking the report by Beagrie et al (2008) as a convenient starting point, and the recommended clauses, sections and other content were tabulated. As each subsequent guide was read a record was made of whether it had similar recommendations, and any different clauses were added to the tabulation. Thus the content of each of the policy guides were compared.

Having established clauses recommended in the literature, the preservation policies were also analysed in a similar manner, by comparing them then tabulating clauses found in the documents. The researcher considered taking an ethnographic approach to the analysis of the policy documents, recognising that rather than assess texts in terms of their correspondence to reality, it is useful to analyse the text to find out how it achieves its effects (Silverman 2006 p. 157). However this type of ethnographic approach would unfortunately be outside the scope of the current research project.

### 3.7 Summary

This chapter has described the underlying philosophy of this research which is based on critical realism, with an interpretative epistemology. It focuses on examining the mechanisms which underpin selection, taking account of the context and individuals' personal experiences. This involves using a qualitative, intensive research design, using a combination of the guidelines given by grounded theory as tools and the very practical guidance from Miles and Huberman (1994) to analyse the data gathered. The research consists of the following elements:

- a review of the literature, including digital preservation policy guides
- interviews with digital preservation experts which orientate the research, provide sensitising concepts and begin to uncover relevant mechanisms
- a series of interviews with library, archive and museum practitioners both in management and curatorial roles to examine their personal views on the factors that affect their selection decision making
- an examination of digital preservation policies
- critical analysis of the above.

The next chapter describes and examines the findings from the series of interviews with digital preservation 'experts'.

## CHAPTER 4 EXPERT FINDINGS

---

This chapter presents the findings from the series of interviews held with digital preservation experts. As the analysis of the interview transcripts progressed it became clear that the themes identified were focused in three particular areas: roles and responsibilities; institutional factors; and conceptual issues. Therefore this chapter is organised under these main headings, then by the themes and issues associated with them.

The objectives of this phase of the research were to gain an overview of issues in selection for digital preservation; to identify stakeholders in the process; begin to identify factors which may influence selection (objectives 2, 3 and 4). The experts will be referred to throughout by the designation 'E' followed by a number. Relevant quotes from the interviews will be rendered in indented italics.

#### 4.1 Professional roles and responsibilities

Different roles were identified by the experts as relevant to digital preservation. E1 drew attention to the question of who is it that decides what is worth spending (potentially) large sums of money on particular objects? She asks:

*'So who decides what is worth that much money? And on which grounds? Is the grounds how many people would consult it? So the crowd is privileged over the one person that might actually discover something vital? Or it's... you know is government going to decide it?'*

This reflects debate over the roles of different stakeholders in digital preservation identified in the literature review (section 2.3.3). According to E1, the solution to these difficulties is that similarly to non-digital records, archivists should stay independent and be a 'neutral third party' in charge of the records. This naturally reflects her archival background. But experts from other backgrounds agreed that selection should be done by professional practitioners acting on behalf of users (as 'proxies'), continuing the roles apparent for traditional selection. E2:

*'It can only be done by proxies ultimately it can only be done by professionals acting in some sense as proxies for a community that doesn't exist or a community that can be invoked but never really tested against I mean in the long term.'*

The experts seem to say that users or potential users of material exist as conceptions of the practitioner and that there is not a direct role for current users in making selection decisions. In the digital preservation literature the role of memory institutions as proxies for stakeholders is seen as unmistakable (BRTF 2010, p.96). However they do recognise that users may have a direct role in preservation by collecting material that is considered insignificant by institutions, which then becomes more valuable, such as video games (E5). This reflects the observation by Lavoie and Dempsey (2004) that 'preservation responsibilities will extend beyond traditional stewards of the scholarly and cultural record'.

Some experts acknowledged that the selection model of 'elites' choosing material to be preserved (similar to that described by Smith (2006, p.51) in a built heritage context) is no longer ideal. There is a need to reflect the needs and diversity of the user community. E8:

*'The main stakeholders have traditionally been the Curator and the Collections Manager.....Now that we are increasingly coming to speak about museums as a nexus around which different cultural viewpoints and interpretations converge, then selection and the right to a voice in the selection and de-selection process become critical elements in representation and diversity.'*

E8 is from a museum background and this issue has been debated for some time in the museum domain. Museums have in the past selected objects from a particular world view and displayed this to users without being representative or inclusive (Marstine 2005, p.9). Selection should now be inclusive of multiple user viewpoints. E8:

*'In principle, selection should be a democratic process, open to multiple voices ... in order to avoid prejudice and misrepresentation. In practice, there is no way to involve every source community in every decision about selection, acquisition, de-accession or disposal.'*

However, E4 questions how users might practically be consulted about selection decisions. Of relevance to the issue of users and their role in selection is the concept from the OAIS model of a designated community (CCSDS 2012, p.1.11). E2 surmised that much work on digital preservation is carried out in the higher education context because there is an obvious community which is easily consulted when necessary. The designated community for the cultural heritage sector may not be as easily consulted as it is very broad. Despite this, the experts expected policies and collecting decisions to be driven by a consideration of user needs. The role of practitioners is to try to guess or understand what users will value. As E6 said:

*'We do these things, and choose what we will apply them to, because we believe we are serving current and future stakeholders, and our collection development and preservation policies try to reflect what we understand or guess they will value about the things we are mandated to collect.'*

The need to serve both current and future users was mentioned regularly by the experts, and E6 also pointed out that the ability of future users to understand and use material may be influenced by what is done to preserve it now.

The experts also saw roles for others within the institution. The experts were clear that the responsibility within institutions does not lie with one individual, as there are many functions which have a role in digital preservation. Responsibility for selection and responsibility for the criteria for selection were differentiated by the experts, as separate functions within the digital preservation process. The latter is clearly seen as a management responsibility. E6:

*'Determination of the grounds for selection should be the responsibility of a corporate group that takes account of as many of the relevant issues as possible, and sets or endorses policy.'*

But the experts argued that there is a reluctance to accept responsibility for making selection decisions. E3 gave *'it's too hard'* as a reason for this reluctance. In addition the experts recognised problems associated with an institution having no one person with overall control of the processes and policies. As E2 says:

*'I think characteristically my experience of institutions has been as far a digital preservation is concerned lots of departments are responsible and none of them take responsibility. If you see what I mean. So you'll find sixteen different departments with some reason to be involved but very seldom do you have a specific member of staff or a specifically properly mandated person with responsibility for it and that's a challenge.'*

This corresponds with the findings from the survey by Verheul (2006) who found that many different departments within national libraries had a role in digital preservation activities, but this is unlikely to be the case in smaller institutions. The issue of management responsibility and cooperation with practitioners is addressed in the literature, where it is also recognised that there is a lack of common understanding and cooperation within institutions (Jones and Semple 2006; Runardotter et al 2011, p.76).

#### **4.1.1 Ethics**

According to E2, practical preservation activities are always underpinned by ethical considerations. He goes on to argue for an ethical outline for practice:

*'What would be nice to have would be a series of kind of principles or expectations and an expectation as to how we work towards this or a feeling of how we work towards this. And also an ethical stance that says we are going to do this and we're going to do this right.'*

However as we do not yet know if the preservation activities that are being undertaken in memory institutions are successful, what works and what does not, it is not yet possible to formulate such principles except in a very general way.

According to E6, selection is an ethical issue where it relates to the interests of other people such as donors, informants, subjects of content, users and other stakeholders; this reflects the view of Lavoie and Dempsey (2004) in considering selection as a social act. E6 suggests examples of potential ethical issues including: not complying with deposit agreements; changing the intention of the creator; not selecting something because someone referred to in it complained; and not complying with collecting agreements with other institutions. Commercial and contractual agreements can drive and influence selection as the institution may be committed to collecting certain kinds of materials or to refer certain types of material to another

organisation (E6). These examples are partly reflected in the consideration of legal influences on selection (section 2.3.5) and are applicable to both non-digital and digital material.

A further ethical issue was raised by E8, who said collecting policies in museums should be more open and democratic to avoid bias in selection, reflecting concern in the heritage literature for the repression of minority voices (Hall 2005, p.26; Samuel 1994, p.211). E8 in particular drew attention to the requirement for institutions, especially museums who may have presented their material from a particular point of view in the past, to be inclusive and representative in their decision making.

#### 4.1.2 Knowledge and skills

The set of professional skills identified by the experts included collection management, risk management, planning, prioritising material and estimating resources. E8 echoes comments by other experts by explaining that many of the skills required are generic information or museum management skills, independent of media:

*'There is a level at which the principle and the skills of digital and physical preservation are identical - planning, resource management, curation, selection, interpretation, familiarity with the legal environment and documentation. Below this is a level of detail at which the particular nature of each different material type demands specific skills and activities.'*

The experts agreed that understanding the goals of the institution in which the activity is carried out is central to successful selection by the practitioner. E4:

*'I say that in an ideal world what we should be doing is looking at the goals and mission of the institution and how we can reflect those goals and the mission of the institution.'*

This is for two suggested reasons. Firstly, if you do not know what your organisation is trying to achieve you will not know what to keep (E3). Secondly, plans for the future could affect selection decisions. For example if there is a plan to increase the size of one department or close another in a university this could affect the choice of materials to collect or to dispose of in the library; it would also affect the decision to preserve that department's research data (E3). This type of contextual knowledge is widely acknowledged as useful in assessing the value of items in the literature (for example Johnson 2009, p.116; Craig 2004, p.49).

So far there is not a 'digital preservation' or 'digital curation' profession; E3 explains this as being because firstly people are not yet temperamentally ready to gravitate toward a role that is 'fuzzy' or includes activities from roles that are currently different, such as IT and practitioners. Some experts rightly identified the problem of institutions not having staff with the

right skills to engage with digital material. This creates problems with selection decision making. E2:

*'The second side of that is an inability to make a decision is where there is simply not skills and simply people have no capacity to make a decision and in those contexts what's happening is the decisions are made or would be made at a kind of really unstructured way.'*

It seems that in this view the lack of training means that selection decisions are being made in an ad-hoc way due to a lack of knowledge and skills. The issue of training has been considered in the literature review (section 2.3.3). The view of the experts seems to underestimate the development of the role in data centres or repositories where practitioners may have subject knowledge and technical knowledge; the Archaeology Data Service (ADS) for example has staff with the title of 'digital archivist' (ADS n.d.). It reflects an assumption that the eventual goal is to have a separate digital curator profession, rather than an institution managing digital material through co-operation between staff with different skills. It also lays the responsibility for the lack of integration of the different skill sets at the door of the practitioners rather than taking account of the influence of management and the institutional environment.

## **4.2 Institutional issues**

The following sections discuss the findings from analysing the expert interviews that relate specifically to the management of the wider institution and their influence on selection.

### **4.2.1 Roles and responsibilities**

The experts agreed that the purpose of selection is very different in libraries, archives and museums. E4 claimed that libraries have '*an easier job*' because most of their material is paper or text based, whereas the problem of preserving photographic material or a museum's material objects is much greater. This seems to underestimate the potential range of material that might be held in libraries. Digital material is still viewed as something 'different'; E8:

*'I devoutly hope that in the next few years, people see past the glamour of 'digital' and come increasingly to think of both physical and digital collections as 'stuff' that needs to be managed, preserved and shared with the public.'*

When asked what the differences or similarities are between libraries, archives and museums, E1 said:

*'I don't think they have much in common other than the difficulties of carrying it out. The methodologies are going to be different; the issues are going to be the same.'*

The traditional approaches to selection outlined in section 2.1 of the literature review does indicate the different traditional aims and practices of libraries, archives and museums, but it seems that digital material is a cross domain issue. The experts discussed the blurring of roles between different types of institutions. A museum may have library or archive holdings for example, or the local authority may have one person in charge of 'collections' across libraries, archives and museums. Two possible reasons were suggested by the experts: E4 attributes this to the spread of digital material; management issues will be the same across domains and there is less emphasis on location (Feather 2006, p.12). E7 attributes the blurring of the role of institutions to competition for funding; collaboration helps to spread the responsibility and cost of preservation (Lunghi et al 2012, p.214). The drive to co-operation may result from both technological and financial concerns.

When focusing particularly on libraries, the principle of selecting for users was highlighted by the experts; libraries have to think of what their current users want to use and manage their collection accordingly, for example by preserving the most used journals or those in demand by academics. E6:

*'Ultimately, by the time we are thinking about preservation decisions, we are (or should be) thinking about what a prospective user could be expected to value about having access to a specific object.'*

The traditional approach in libraries is very user-focused (for example see Clayton and Gorman (2001); Harvey (2005, p.58); and Johnson (2009, p.108)) and this approach also applies to selecting digital material.

Also in libraries digital material such as e-journals may operate under different business models than traditional material, which affects collection management and preservation, as discussed in section 2.3.3. E5 explained:

*'I guess the big difference there is that electronic material tends to be licensed particularly the journals so you have an issue around business models that doesn't really exist with traditional or at least in the same way. Because you're renting a copy the whole issue of how you then guarantee access is different because you don't have it directly under your own control.'*

This may lead to the need for third party services to provide that guarantee should the publisher go out of business. It may also change the relationship and provide complex contractual arrangements between libraries and publishers. National licensing schemes such as the NESLi2 licence from JISC (JISC n.d.a) aims to help libraries with this issues.

Archives are seen as having a *'more pressing need'* (E2) to engage with digital preservation earlier than other domains. E1 said that earlier engagement in archives is due to the greater consequences for record managers of not preserving necessary records due to legal requirements. The aspect of archival appraisal which is very different to the approach to selection in libraries and museums is that there is an assumption that material will be disposed of. This was highlighted by E8:

*'The archival approach to selection is fascinating - both because of the very important principle of acquisition on the presumption of disposal (yes, we'll take it, but we're making no promises that we'll keep it) and because of the commitment to filtering out the ephemeral to create a lasting social, historic and financial record.'*

This view conflates the traditional roles of records managers and archivists, acknowledging the role of the archivist in creating the archive. Selection of digital records starts when the records are created and if records are regularly destroyed in the usual course of business then those with a permanent importance will be kept through these processes. This is encapsulated in standards such as BS ISO 15489-1:2001, which applies to records in all formats. Accordingly, appraisal should be part of the ordinary business process in an office and be embedded in the appropriate systems. But records can be created in an ad-hoc manner and not organised according to standards which makes appraisal more difficult; E1 uses the example of artists:

*'But when you're dealing with artists, literary artists or visual artists even worse you... what do you appraise? You appraise the importance of the artist because you don't have a clue.... you have to make an assumption that no matter what, it will be important to have the stuff of this person no matter what the stuff is.'*

This is a reflection of the process of macro-appraisal, in which the value of items is determined by reference to their larger function rather than on an individual basis (Shepherd and Yeo 2003, p.151).

E2 thought that in museums digital material is seen as a secondary or subsidiary collection, so curators will rather concentrate on what they see as their main collections of accessioned objects which they are funded for. E8 has a museum background and was therefore very illuminating in describing the problems of digital material in museums. He described the selection process in museums as expansive rather than reductive, echoing the traditional form of selection described in section 2.1.3. According to E8, museums have always had an acquisitive urge to collect without any clear idea of what they are going to do with the material once they have it:

*'In a ferment of exploration and discovery, the curators of that generation said 'yes' to a huge quantity of physical material, and in their enthusiasm came to see the more formal elements of Collections Management as a chore, to be put off until tomorrow.'*

It would seem that this emphasis on collecting before considering how the collection will be managed is a historical one in museums. This means that museums do not have the appropriate infrastructure in place to manage their digital material in more than the short term.

E2 said:

*'We have these really clever really inspirational programmes engaging people with digital technologies. Led in some senses by the museums sector. Without the necessary infrastructure or wherewithal to sustain that.'*

The lack of infrastructure in the museum domain to engage fully with digital preservation is echoed by E8:

*'So, in the same way as the physical artefacts which we have neither space, nor people, nor money to look after, we now have a generation of digital material with insufficient funds, time or infrastructure to preserve it.'*

E8 pointed out that the national museums do not show the same sort of leadership as the BL and TNA for libraries and archives, so the situation in the museums domain is more fragmented.

The experts also pointed out that national institutions have a particularly demanding role in comparison to local institutions. They have a large and diverse user community (the nation), a long term preservation need and they must preserve large amounts of both paper and digital material. E5 termed the latter a 'dual mandate' for the BL, without concomitant dual resources. A particular thread in the expert responses to the question of responsibility is that the responsibility of the national institutions has not changed - they must collect what they have always collected; the preservation imperatives of national institutions are broad based and long term. This can clearly be seen in the responsibility of the BL and other national libraries for preserving legal deposit material (see section 2.3.5). Smaller institutions have responsibility for a smaller, better defined area. National institutions have statutory responsibilities which they must discharge. The experts felt that ultimately it will be major institutions that will have the responsibility for preserving and selecting.

The experts acknowledged that each institution has a responsibility to understand its own needs and make a decision as to whether they have the inclination or capability to undertake preservation. The priority for the institution needs to be determined and be clear in policy where the responsibility lies. E3:

*'I think each institution has a responsibility to understand what their information needs are; they may well come to the conclusion that they don't have the inclination or the capability or the capacity to undertake preservation or to take it on board as one of their institutional problems. And I think that's a valid response.'*

Institutions also have an ethical responsibility to other stakeholders e.g. donors, subjects of content and also to the public. E3 described a hypothetical situation in which a smaller institution may not have the resources to preserve a particular collection of high importance and it is their responsibility to know when to hand the collection to another institution that is able to. E3 was here referring to the report from the Blue Ribbon Task Force (2010) which examined economic factors in the sustainability of digital preservation. The report includes many mentions of such 'handoffs', mainly in terms of when the value to an individual or organisation of digital assets is not enough to sustain their investment in preserving them.

Museums have the problem of who should pay for digital preservation; there is no one whose interests are served by paying for preservation (E8); there may be a lack of incentives for the institution to preserve digital material, similar to the issue described by Currall and McKinney (2006). The experts identify some benefits to an institution of successfully engaging with digital preservation, including that it is able to demonstrate that it can curate material effectively, so enhancing its reputation. E3 said:

*'I think perhaps a more compelling sort of argument for selection in my mind would simply be going back to this notion that that your organisation is demonstrably good at handling information.'*

Also included by the experts are agreements with donors which may determine what happens to their material, publication programmes, planned or future exhibition programmes and other priorities arising from institutional activities. Also the institution may be able to derive additional benefits from the material it has selected to preserve, such as being able to charge for access or receive a fee for allowing it to be made available. However the tangible benefit or output of preserving digital material is often difficult to determine as it may be unseen, hence institutions are less likely to take responsibility for the material (E8).

#### **4.2.2 Risks**

Two aspects of risk were raised by the experts: risk to the material and risk to the institution or organisation. Some experts suggested that a driver to selection was the need to preserve material as the medium it is on is deteriorating, such as non-digital material that is considered for digitisation to reduce handling the originals. E4 was particularly insistent on this:

*'I'd want to again press the button around the issue of physical fragility and I think there is selection that is being done and quite rightly so for digital preservation on the basis of physical fragility of traditional materials.'*

E4 uses 'digital preservation' to refer to digitisation to preserve physical material, reflecting the dual meaning of the term found in the literature (Conway 2010, p.64-65).

E6 explained succinctly that a difference between analogue and digital material is that:

*'The nature, mechanisms and life-cycle of deterioration and loss are largely different (although there are some parallels such as the physical deterioration of digital carriers like discs and tapes)'*

The experts agreed that the nature of digital material means that selection decisions must be taken quickly; those faced with selecting digital material cannot allow 'time' to select what is most valuable. Digital material has a 'cliff of loss' (E6) unlike traditional material which has a more gradual process of degradation; there is no way to 'see' if intervention is necessary to prevent loss. Due to this need to act more quickly, digital material needs someone to take responsibility for it to avoid unplanned loss and processes need to be put in place to check the ongoing quality of material (Harvey and Thompson 2010); it is very expensive to replace. E4 described this quality of digital material as 'nowness':

*'And whereas with the digital we've got a now to the collection if you don't collect now chances are if you come back in five years time it just won't be there or it won't be in a form we can use or it will be un-understandable it will be a database without any record as to what the fields mean stuff like this.'*

This refers to the threat from obsolescence and also the risk of not gathering and preserving contextual data (as described in section 2.2.1). In addition to drawing attention to technical risks to material, the experts raised the theme of risk often in terms of the organisation. They regard it as an institutional decision as to how much risk it is prepared to bear and who is responsible for owning that risk. It was suggested that institutions keep material for 'insurance' purposes, to guard against anything going wrong in the future. E3 again:

*'I mean there's the kind of old insurance argument as well particularly in pharmaceutical companies or aerospace companies to keep those CAD drawings because your planes are going to be in the air for fifty years and you don't know whether they're going to crash or not.'*

According to the experts, it is more likely that the responsibility for risk management (and digital preservation) is dispersed through many departments, for example collections management, reprographics or audio-visual departments. E2 claimed that publically owned institutions may be more risk-averse than other types of institution.

Related to the issue of risk is that of the trustworthiness of institutional repositories and the trustworthiness of external companies or institutions that provide infrastructure or digital preservation services. The experts highlighted the centrality of trust in the development of centralised or outsourced digital preservation infrastructure. E2:

*'So there's certainly scope for shared infrastructure there's no question about that not just locally, not just nationally but internationally. But it has to be done with appropriate cognisance of trust issues....'*

Bearman (2007) suggested that cloud computing could provide a solution to selection, as a copy of material that everyone requires could be stored and access could then be provided, but with regard to this E2 said *'I think that the technology is ahead of the politics'*. This he explained as being due to problems such as data protection, access control mechanisms, and personal security. The problem is that there is no method of measuring the 'trustworthiness' of companies that enables confident decision making. E2 again:

*'The only way we would know we could trust Amazon would be if we could somehow go up and measure them and compare them against a predefined set of criteria which we'd all agreed and we don't yet have that set of criteria so until we have the trust metrics we are going to struggle to make, to be able to trust each other in some sense.....'*

Methods of assessing risk have been created, such as DRAMBORA and TRAC (section 2.3.5), but these are only applicable to repositories. Much reassurance is needed by practitioners and managers; the question of why should we trust the cloud and related services has not yet been answered as the company could stop the service at any time (E5). Additionally, E3 linked sustainability of organisations to trust:

*'There's a sort of it's almost a core of you know a) do we trust these digital materials to last well that depends on how trustworthy your organisation is and how sustainable it is as sustainability is largely about trust.'*

The BRTF report suggests that to engender trust, third party organisations need to be open and accountable with clear agreements, perhaps with legally binding, detailing processes and outcomes (2010, p.43). Despite identifying problems of trust, the experts were positive about the potential of collaboration for the future between institutions in particular. The benefits of

collaboration in infrastructure for smaller institutions were noted (E5), where the curatorial function could stay with the institution but the IT infrastructure could be shared. Other functions could also be shared; E2 mentioned the example of the PRONOM file format registry held centrally by the National Archives and also the development in Denmark of a national storage infrastructure.

### 4.2.3 Legal issues

Many legal issues were raised by all the experts as being particularly important in relation to digital preservation and selection; there was agreement that the legislative framework influences selection decisions and digital preservation activity in both overt and covert ways. Intellectual property rights may be infringed with the need in the migration process to change the format of an object to make copies. Legal deposit is also an important issue; E4 felt that publishers have been lobbying very heavily against electronic legal deposit especially in the newspaper industry, as they think the BL is straying into their territory and this may affect their commercial viability and monopoly. The effect of the new regulations, which represent a compromise between the different stakeholders including publishers, are still to be seen but it is clear that the concerns of publishers are reflected in them. Restrictions on where and how many people at once may access the material (s.23 allows only one computer terminal is available to readers to access the material) are ascribed by E3 to the influence of publishers and claims that:

*'It's like... and they've said in the act and they make no bones about it that we are sticking as close as we possibly can to a print paradigm with this act so forget all of the you know don't worry about web thinking.'*

It seems that the relationship between libraries and publishers may be more complex in the digital environment due to a perceived intransigence of publishers to change. This is understandable as they are concerned with protecting their intellectual property but it limits the potential for use of the material.

A further legal issue discussed by the experts focused on privacy of confidential data in personal collections, as the subject or their family may not want certain items preserved. E5:

*'Personal confidential data comes in particularly when it comes in personal collections what you do about the love letters to the mistress [laughter] all those sorts of things I suppose.'*

Certain items may be embarrassing or damaging to someone's reputation. Issues such as this in risk-averse institutions could influence the activities that the institution chooses to undertake in order to avoid complex legal situations, especially where material relates to third parties.

Little research has been done into preserving personal collections; an exception is the PARADIGM project (2008). Legal considerations examined in this project include compliance with the Data Protection Act (1998) which may be particularly relevant to digital archives as the people referred to in them may still be alive (PARADIGM 2008). The approach suggested by PARADIGM (2008) relates issues of confidentiality not to what is selected, but to what is made available and when. The project was careful to focus on records of the subject's professional life, noting that participants were reluctant to place personal material or that relating to third parties in a library, reflecting the observation made by E5.

#### 4.2.4 Policies

Surveys of cultural heritage institutions which ask about preservation policies reveal that many organisations have not yet devised a policy (Waller and Sharpe 2006, p.16; Beagrie, Rettberg and Williams 2008, p.1 for example). E3 mentioned that policies are still relatively rare:

*'At the moment I don't think that's particularly widespread. I think that people have some of these other ones [policies regarding information governance] because they're either obliged to or it's a good idea to do so but I'm not sure that there's a great many that have specifically said we are now going to talk about preservation in terms of how are we going to sustain stuff we are going to need.'*

The experts suggested that ultimate responsibility for policies should lie at a very high corporate level in an institution, perhaps at board level, for example with a chief information officer (E2). Changes to policy should lie in a group that can take into account all the relevant issues (E6). The role of this group or individual includes responsibility for managing risk and ensuring the institution complies with statutory requirements. The policies themselves may need to be very high level and give an overview of what it is the institution wants to keep (E2); the policy will then drive other activities (E3) and justify strategic decisions to both internal and external audiences (E7).

Some experts said that collecting policies are the same for traditional and digital material and what drives digital preservation policy is the same as for traditional material - what is collected is not necessarily going to change, especially for the big national institutions as their remit is no different (E2). According to E2, smaller institutions can define the scope of a collection policy more easily than larger organisations, even if they don't have the resources yet to deal with digital material. This was suggested by TNA (2011b, p.6), in recognition of the reputational value of such a policy. Depending on institutional mission (and where the institution sees its preservation needs) there may not be a need for an in depth preservation policy; it may be that a reference to an external preservation specialist service, a repository for example, is sufficient. This is recommended by Beagrie et al (2008, p.27) to frame service level agreements with third party services.

#### 4.2.5 Finance and costs

E2 raised the question of whether selection would cost too much to do:

*'The key question for me is whether, still whether, selection actually is economically viable so are we actually going to do it? Or is it cheaper just to keep the stuff and we'll invent some clever mechanism to do it?'*

The literature seems to indicate that a selective approach is currently considered the cheaper alternative (Paradigm 2008; Whyte and Wilson 2010). The experts agreed that costs are one of the most important drivers to selection and in selection decision making, despite the general assumption that digital is cheaper than paper. The experts suggested that cost is often an unstated criterion, or driver, for selection. E4:

*'And there is a fallacy that digital is free, that storage is free and that what you will do is that you will replace your physical activity... your activities relating to the physical traditional collections with the digital... but the reality is that there isn't enough money. And that drives an awful lot of selection decisions. Finance. Where the money's coming from and how much you can afford to do.'*

It is clear from the literature that funding models have an influence on what is kept (BRTF 2010, p.58). The experts suggested that the finance associated with specific project funding would drive selection, as the conditions attached to that money would determine the criteria for that project. It is the existence of project finance that drives selection rather than the needs of the institution. Rather cynically, E4 suggested that:

*'It's that we [...] money from wherever we can get it. Once we've got the money we plan how we're going to spend it. We then decide that our information goals then magically align themselves with the project we're going to do with the money that's available which obviously will have a huge benefit to our community because we said it will.'*

Money may not only be associated with a particular project, but it may be 'ring-fenced' by institutions for certain activities only (E5). The problem of how much an institution can afford is affected by the cost of various factors. The experts discussed the costs of storage in particular. Paper-based storage is becoming more expensive and this may drive institutions toward selecting digital versions, for example the increasing prevalence of e-journals in academic libraries (E5). However the cost of storage for digital material is potentially a problem, more specifically the cost of maintaining the storage and the expertise needed. The costs of storage and an overview of research in the digital preservation field is described by Rosenthal et al (2012); they conclude that the rate of decrease in the cost of storage seen in the past will be much slower and hence storage will become a more important cost consideration (Rosenthal

2012, p.7). E4 suggested that storage costs may affect the choice of format; choosing those, such as jpeg2000, saves storage space rather than being the better file format.

### 4.3 Conceptual issues

The final three issues identified through the interviews with digital preservation experts are conceptual.

#### 4.3.1 Significant properties

As seen in section 2.2.2., the concept of significant properties is imprecisely defined yet prevalent in the digital preservation literature (Dappert and Farquhar 2009), so it is unsurprising that the experts discussed this concept. The type of material under consideration affects the ease by which significant properties can be identified, for example the significant properties of a data set are much easier to determine than digital art. E2 explained that from his point of view 'significant properties' is a useful concept as it allows us to measure whether our preservation actions have been successful:

*'what you can say is that we've said that these are the things we think are really important we've defined that explicitly at the start we've undertaken a number of preservation actions for that so emulation or whatever. At the end of that process are the significant properties still present, recognisable, authentic?'*

This closely reflects the reasons for assessing significant properties found in the literature (Wilson 2007, p.7).

E6 questioned the use of the word 'significant' by asking '*significant to whom?*' This expert made it clear that different stakeholders will have different views on significant properties:

*'We have tended to move away from that term in [...] because we have had so many arguments about "significant to whom". Our IT people tend to think about significant properties in terms of technical properties, so we often find ourselves going around in circles on the question of whether any change in any object is acceptable...'*

On a broader level E2 made a similar point that the properties of digital material which are significant in a museum context will be very different to those in a library or archive context. Use of the 'preservation intent' concept could be useful here as it includes a broader range of characteristics than implied by 'significant properties' (del Pozo, Stawowczyk Long and Pearson 2010, p.295-296).

### 4.3.2 Value and criteria

E1 points out that digital preservation can be expensive, such as to develop emulators, but the issue is *'...it's not just the money it's what is worth the money'* i.e. what is valuable enough to spend large amounts of money to save? Value is a complex issue. In archival terms, appraisal is a decision on value. E1, from an archive background, took the view that value is not the right term for appraising records as each record is equally important to the meaning of others. Value is in this view highly contextual. She says:

*'And the grounds on which the decision on value is made has not changed and will never change. Because value is a relative concept and it very clearly depends, that the assessment of value very clearly depends on the mission or mandate of the repository, on the specific context, cultural context, legal context... What has changed enormously is the process for actually conducting appraisal.'*

This reflects value as described by Graeber (2001, p.40) as *'simply meaning: giving value to something is a matter of defining it by placing it in some broader set of conceptual categories'* or in other words *'value is meaningful difference'*. However the problem that Graeber (2001, p.43) identifies with value as defined here is that it is not evaluative, in that we cannot say by how much something is more valuable than something else. A measure of relative value is necessary as it may not be possible to afford to save everything that meets the mission of the institution; Atkinson (1993 p.98) identified this problem when he stated that after material is ranked for selection purposes, the question is then an economic one.

E6 differentiates value in two ways, addressing both intrinsic, meaningful, value and extrinsic value:

*'The first is the concept that all items in a collection carry values for which they were collected. These are the intrinsic characteristics they have which make them meaningful to users. The second concept involves comparative assessment of the significance, importance, even monetary value of items.'*

The second aspect of the concept of value helps to explain the importance of economic value because it is easy to make comparisons between items in economic terms. Formal systems for determining value include ranking scales which determine relative preservation priority, as described by E7. E4 suggested the measurement of value in terms of the impact that the material may have; the funding of selection could be evaluated on the basis of the impact it will achieve. Whilst this has potential and some work by Tanner (2012) in developing ways of using impact has been done, it has not yet been explored specifically for selection.

Making decisions on value for digital preservation purposes may mean trying to take into account many types of value and a number of different types were identified by the experts in

addition to economic value and impact. Cumulative or contextual value is not only relevant in an archival setting (see section 2.1.1) but also in a library; E5 explains the importance of collection building in libraries:

*'Actually I suppose another thing is around collection development in terms of selection. It's added value. A single item may not have value on its own or have limited value but as a group it may be important so there's this whole idea of building collections in particular areas and the cumulative value.'*

Deferred (or potential) value is described by E2 as perhaps the most intangible type of value; it is often least related to policy goals and may bring in little obvious return on investment. E2 links this strongly to risk management, where the value of material to organisations is for potential future evidentiary purposes. Keeping material allows an organisation to comply with statutory or regulatory requirements. E2 also described what he referred to as 'opportunity' value:

*'So when it is very expensive to replace something that we have inadvertently lost and which we suddenly realise we need....you know where NASA lost the data from the Viking landers and its very expensive to land things on other planets and you don't want to have to do it again. So keeping the data from that is seen as an advantage and there was it was possible to reconstruct the data but only after significant effort. So a value there is if you like an opportunity value or an opportunity cost.'*

The term 'opportunity value' is misleading, but the concept of choosing to keep material because it is too expensive to produce again is reflected in a criterion suggested by Whyte and Wilson (2010). 'Non-replicability' is suggested, by which they refer to the question of 'Is *the cost of replicating or re-measuring the data financially viable?*' In the context of preserving research data, such as the instance referred to by E2 with the Viking Landers, it is an obvious question, but it is also applicable to other types of material which are unique. A second type of deferred value identified by E2 is potential economic value. E2 used the example of the music industry, which will derive value from its future back catalogue and the exploitation of the intellectual property rights it holds. Deferred value to the institution is described by E3 as the value of preserving material ethically and professionally over time. This leads to two potential benefits - an institution becomes known for good curatorial practice and it may enhance its teaching or research value through holding unique collections which have been kept safe.

Value can also be considered as cultural value, similar to many of those described by Throsby (2001). The experts considered 'heritage value'. This type of value is related to the way in which objects 'bear witness' to historic events, whilst possibly having very little economic value. An example given by E2 was of a set of drawings brought into his previous workplace by an elderly lady. She had drawn them as a child when interned in the Nazi concentration camps, and

described the things she had seen. Many of these types of items may have specific local or personal interest and the justification for keeping this material is more difficult to state for an institution.

Criteria are tools for assessing value. The experts identified a range of criteria relating to selection which they felt were particularly important, such as the mission of the institution, value, whether the item is catalogued, and the condition of the item. E7 mentioned commercial considerations with the possibility of gaining 'additionality' which he explained as the possibility of gaining fees for making their material available through third parties. Whilst commercial potential as a criterion was not found to be highly important by Gould and Ebdon (1999, p.12), commercial partnerships were identified by Moore, Jeffery and Richards (2010, p.53) as one of a number of potential funding models, so this may become a more important consideration.

The experts' discussed using the format of digital material as a criterion for selection. The problem with this is made clear by E2:

*'The decisions would be made structured around some of the things like media, criteria which aren't actually really the criteria which we ought to at an intellectual level be content with but which work. So save everything that's on a CD but forget the stuff that's on obscure floppy disc.'*

According to E2, whilst format should not be a criterion on an intellectual level, in practice it is because some formats are easier and cheaper to preserve than others. This echoes the point made in the literature review (section 2.3.1) that the technological ability of the institution are important criteria for digital preservation and if something is difficult to preserve then it may not be selected. It is also suggested as a criterion by practitioners (section 5.4.2).

It is clear from the literature review that value is not a static characteristic but can change. The view that a decision can be made about how long the 'lifetime' of an item should be before it is deleted was put forward by E8 who works within the museums sector; this is effectively a decision on how long an item might be valuable for. This reflects a records management approach where records are destroyed if no longer valuable. E8 says that:

*'One really interesting new discussion that has emerged from the museum community is that maybe things have a lifetime, and that sometimes we can make decisions about how long that lifetime should be.'*

E8 goes on to say that not to select could be seen as a kind of 'stealth deaccessioning'. As some traditional material is physically very fragile, selection for digitisation is done on the basis of condition as there is a huge risk to physical collections. If the materials are not selected then

this is effectively deciding they will not survive. This use of negative criteria is not widely discussed in the literature, although is apparent later in the interviews with practitioners regarding criteria (section 5.4.2).

### 4.3.3 The future

The concept of the 'future' was used in two ways by the experts. Firstly they considered the 'future user'. The role of the practitioner is to 'guess' what these unknown future users may want or need. The question of what stakeholders over time will value is central to evaluating material.

E6:

*'We do these things, and choose what we will apply them to, because we believe we are serving current and future stakeholders.'*

E2 made the fundamental point that it is not possible to predict what future users will need or want, and that we do not know if digital preservation has been a success until those future users interact with the material:

*'We don't yet know what success looks like and until we do we will continue to spend more money on it than we probably need because we don't know if it has worked yet. But it is like a relay race - it will never be finished, it will just be as it is today. Success will be if future users can exploit and use the material but we are not at that point yet. It is difficult to tell how users will interact with material for example climate change data from ships logs. It is not possible to predict what will be important in fifty years time.'*

Experts identified a range of issues that may become important in the future. E7 suggested traditional conservation work on books and paper will become less as there will be a more minimalist approach, by digitising then storing the original. This she attributed to a change to people wanting to access material online rather than in hard copy; in her experience it is also cheaper to digitise a book and store it than perform specialist conservation work on it. This may be an optimistic view, as it has been suggested that digitisation may increase demand for the original (Peach and Foster 2013, p.15).

Many of the experts suggested that financial considerations will become the decisive factor, and more will have to be done with fewer staff on increasingly complex material with greater user expectations and less resources. As there is no obvious funder for digital preservation activities, within museums especially (E8), and no clear business model to provide for it, we are likely to see a widespread loss of the digital record, according to E8. Most museums may be more concerned with their own survival than taking on long term responsibility for preserving material that will incur potentially high future costs. In order to combat this, public projects should be thought through on a sustainable basis from the beginning. Digital preservation needs to be incorporated into an ongoing management process, rather than treated as a project, to be

sustainable (BRTF 2010, p. 107). However deleting digital material seems easier than deaccessioning physical material and so there may be a temptation to keep physical material rather than digital; as E8 suggested:

*'Pressing 'delete' or reformatting a hard disk is \*much\* less conspicuous than a skip parked outside the museum [respondent's emphasis]'*

The potential for responsibility for different aspects of digital preservation to be shared was recognised by some of the experts, who suggested that advisory services such as the DCC may become more important in disseminating services and tools and in providing training and advice. Third party services such as data repositories providing storage may also become more important. However, the technology is ahead of the politics as there are not the appropriate levels of trust, security, legal frameworks and others requirements (E3). E3 also mentioned that an issue which may develop is that of environmental impact - digital preservation is not carbon neutral and the more material we try to preserve to greater our environmental impact with the extra servers and equipment needed. So reducing the amount of data you have responsibility for may be both an economic and a 'green' issue, especially if environmental legislation begins to affect organisations. This issue is not addressed yet in the literature.

#### **4.4 Conclusions**

The themes and issues identified through the interviews with digital preservation experts relate to influential factors in selection. The experts addressed a wide range of issues, building on those identified through the literature review. The factors influencing selection identified by the experts were: roles and responsibilities; ethics; knowledge and skills; risks; legal issues; policies; finance and costs; significant properties; value; and aspects of the future. These are taken forward in the questions asked of the practitioners (see appendix 4 for the practitioner interview schedule). It is clear that although the experts are not all archivists, there is an archival theme underpinning their view of many of the issues raised. Ethics for example were considered as two issues: compliance with legal requirements and agreements or the need to be neutral in collecting.

The experts had a broad overview and they looked forward, identifying possible issues. They highlighted value as an influence. The types of value identified by the experts related to current value; contextual value and cumulative value which derive from being part of a collection and heritage value, where items are 'witnesses' to the past. A further type of value identified was deferred value, or potential benefits the material might bring in the future, such as future economic value from exploiting rights, reputation value from being a good custodian or future research value. Deferred value is not reflected in the criteria suggested in the literature (section 2.3.1) which focus more closely on current issues and abilities of the institution.

The idea of the future was related by experts both to practical issues such as costs, and also to future users. The experts highlighted the user in discussions on value, roles, and designated communities. The latter concept was recognised as problematic, as experts described how different stakeholders will have different views of what is significant. There was a suggestion that different user groups should be better represented in selection decisions and they may have wider roles than merely as consumers however they did not suggest an alternative method of representing users and the assumption that practitioners should continue to be responsible for acting as a 'proxy' or on behalf of the user was apparent. The experts also identified a number of further stakeholders in selection and preservation, including senior managers, risk managers and IT personnel. They reflected findings from the literature review by highlighting the problem of having no one take overall responsibility within an institution. Similarly to in the literature, the experts felt that digital preservation issues are a cross-domain responsibility.

Another theme which was apparent in the expert interviews is that little will change with digital material; E8 for instance hoped that practitioners will eventually see digital material as 'stuff'. The experts suggested many of the skills that are useful for managing preservation and selection are generic, so practitioners already have many of the skills that they need to select for digital preservation. The roles of setting criteria and applying criteria were separated by the experts, demonstrating that selection should be closely tied to the mission and policy of the institution determined by senior management, similarly to non-digital selection. According to the experts little change is seen in what will be collected, it will just be in a different format, though this has not been examined in practice.

Despite the emphasis on similarities between selecting and managing non-digital and digital material, differences were identifiable through the interviews and these focused on skills, material properties and criteria. There was an assumption that the development of a digital preservation profession was desirable which would have both curatorial and technical skills specific to digital material. It is not clear at this point of the research whether this is developing or whether it is desirable in all contexts. The rate of loss was also different to non-digital material, and this was emphasised along with the issue of risk to the material. Format was recognised by the experts as having an influence in terms of the criteria for selection, where it should not be a criterion but in practice it is. Choosing formats has implications for future costs and the experts were clear that this is an 'unstated' criterion. Other unstated criteria included the cost of storage and the organisation of budgets within institutions. This suggests that there may be other unstated criteria and this is further examined in the interviews with practitioners (see section 5.4.2). The differences identified by experts were reflected closely in the literature, which is unsurprising as some of the experts have published widely on digital preservation.

The influential factors identified in this section of the research were then taken forward into the interviews with the practitioners; findings from these are reported in the next two chapters.

## CHAPTER 5 PRACTITIONER INTERVIEW FINDINGS

---

This chapter relates the findings from an analysis of the interviews with the practitioners. Similarly to the experts, the practitioners will be referred to by the designation P then a number, as detailed in section 3.5.1, and quotations from the interviews are shown in indented italics. In this chapter the findings will contribute to the fulfilment of objectives 2, 3 and 4.

## 5.1 Roles and responsibilities of the professional

The practitioners were asked questions regarding their own role in selection.

### 5.1.1 Selection as a professional role

Selection is seen as a professional function:

*'But yes anyone making selection decisions would have to be professionally qualified or ought to be professionally qualified' (P7)*

P7 states that professionals have a better idea of the future and the needs of future users than others who consider only current needs. This reflects assumptions found in the literature (Shen et al 2011, p.216) and the opinion of the experts (section 4.1). Consideration of professionalism leads P12 to question the use of volunteers in archives. Volunteers have not had professional training, so in her view they are not able to be dispassionate about the material. P12 implies that they might therefore allow their personal preferences colour their selection decisions, undermining the role of the archivist as a collector not creator of archives (Duranti 1994, p.343). This seems short sighted and dismissive as volunteers may be professionals themselves. Both P12 and P7 earlier claim specialist abilities and knowledge for archive professionals that others without their training do not have. They feel that those with professional training can exercise better informed judgements and have procedures in place that mean that material is considered consistently and fairly. Other archive respondents mentioned adhering to further archival principles, including demonstrating authenticity, integrity and being accountable. It seems that practitioners are keen to differentiate themselves from others, emphasising their continuing relevance and special nature. This excludes firmly the possibility of future collaborations with users; users are circumscribed by the desire of the professional to be seen as having special knowledge or expertise.

### 5.1.2 Determining value

The respondents were asked about the concept of value in selection and how they determine it when selecting material. Respondents from all domains mentioned financial value, but they tended to deny interest in the financial value of items, showing reluctance to value items in this way. Financial value is considered a *'bit of an irrelevance'* (P21 and P22; P19). However ways of valuing items financially were mentioned by respondents. For digitised data there is the cost of recreating the digital resource from the analogue originals (P1); this is similar to the 'opportunity cost' mentioned by E2, and is suggested in the DPC Handbook (2008, p.35). The financial value of non-digital items can be set by 'sale room precedent' or the actual value on the open market (P9). In addition, P9 identifies a financial value for photographs in particular in an image which has a commercial reproduction value; this recalls the 'additionality', suggested

by E7 that could be gained from commercial partnerships (Moore, Jeffery and Richards 2010, p.53).

Archivists were keen to differentiate between financial value and archival value. Archival value lies not in the medium (though some items may have this) but instead is related to the relationship it has with other items (Doom 2004) and its value as a record of something (Schellenberg 1956, p.139). The item must fulfil certain conditions to have value – be of local significance to a local authority archive or be a non-current record for example. In their descriptions of value the archivist respondents demonstrated their use of the theoretical archival value found in the literature. Other respondents were less clear on how value could be measured. Value is a relative characteristic; P5, a curator, described value as not a '*black and white*' decision but as a '*spectrum*'. This clearly allows her to use her own judgement in deciding value. Value may be 'measured' in relation to other factors, such as by how well it fits with the collection policy (P13); how well it fits with the mission of the institution; or, according to a museum manager, against collections or items held in other institutions, where an object that is not collected elsewhere will have greater value than something that is. Few respondents mentioned any formal or external methods for deciding value. Only one library manager stated that her institution was looking at using a significance framework developed externally. However they also described 'guides' that aid them in making decisions. The individual practitioner, especially librarian respondents, described consulting other experienced staff in the institution such as specialist acquisition staff, or subject experts from outside the institution (see 5.2.2).

Respondents valued material more highly if it is used or fulfils a particular function. Examples suggested by respondents included material that:

- contains information about area of responsibility
- records something happening
- satisfies a user need
- fills a gap in the current collection
- is a key item in a story
- is a representative example
- adds to an existing area of interest
- says something new about a famous person or his/her context.

Use value is not just contemporary use but also potential use in the future; an item can become more significant when a use is found for it. P20 made the distinction between value and 'worth', where value to her is financial but worth is whether it may be of use in some way. Potential worth is related to potential use, such as in material where restrictions due to copyright will end at some point and the institution will be able to exploit the material. Because of its potential worth it should be collected (P20). Practitioners also acknowledge they can influence value

directly by finding new ways to use material which will increase significance. P10 relates a similar collecting justification in his institution when he says that they collect material from less well-known figures, not just famous people, as it has potential value for future historians or other researchers. Respondents mentioned that value is related to the audience for the material; if they find it interesting then it is more valuable. The subjective nature of value is seen as extending to the user; it is different to everyone who wants to use material. This makes assessment of value more difficult for practitioners as they are trying to guess the values that other people may have and these are unknown, especially for future users. P14 describes this dilemma as:

*'What is fashionable history, what is fashionable today may not be in the future and there are other things that we don't understand now that become of interest in the future...'* (P14)

Historical value was mentioned by many respondents, but most especially by archivists. Historical value is partly based on the significance of the thing described (P15) and partly on whether it adds to the representation or understanding of it (P7). Archive respondents in particular felt they had a responsibility to 'keep history safe' meaning that they feel they are custodians of evidence about 'the past'. The archivists still have a sense of themselves as being closely related to historians and in part recognise their role in producing history by choosing what to keep (P7). They also recognised that archivists try to be objective and not to 'skew' the representation of history, reflecting the traditional view of the archive role (Jenkinson 1937; Duranti 1994, p.343) but that this is difficult:

*'...you can't do too much in the way of thinking about it and saying oh well in 50 years' time this is what it's going to look like so I'll collect that because this is what it's going to do. So it's really hard you kind of have to look at what is going along now and it's almost like looking to the future but using now as your marker if that makes sense?'* (P12)

This according to the literature and the experts this seems to be an issue to which there is no one answer, but the archivist clearly makes at least partially subjective appraisal decisions (Reed 2009, p.124; Cook 2011, p.177). Respondents from libraries did not focus on history as much as the other domains. Family or local history researchers are one of their user groups, especially for the local studies librarian, but the emphasis is on providing them with information rather than 'keeping history safe', which was more a preoccupation of archive and museum respondents. Many of the libraries had historical collections but these were often considered archive collections and so were cared for by archivists working in the library.

Respondents recognise that material is easier to value with 'hindsight', after a significant period of time has passed. Some practitioners struggled with choosing contemporary material that will

become historical heritage material. A question about heritage was not directly asked, but when the issue arose they were reluctant to consider outcomes in terms of preserving heritage material that will be passed on to the future. When the issue was raised with P16, her first response was that it is 'scary', and that:

*'Well no, if I thought like that on a day to day level... you'd probably not choose very much, you'd get so hung up about it that you could get stuck quite easily.'*

In order for practitioners to select material, they alluded to what P5 refers to as 'historical sense', a skill which comes through professional training. This helps the practitioner evaluate contemporary material, looking forward to estimate what will be important to historians in the future. She goes on to say that it is a very difficult thing to do and admits it has not always been done well. This is echoed by other practitioners, especially archivists. P10 thoughtfully states that:

*'You don't want your sense of value to be too rooted in the particular moment'*

Respondents were clear that value is decided on by professionals and depends on judgement, developed through training and experience.

### **5.1.3 Subjectivity**

In order to demonstrate how they fulfilled their roles and responsibilities, practitioners were keen to highlight aspects of their professionalism and expertise. Using professional judgement seems central to the conceptualisation of 'a professional' for respondents in all domains. The practitioners in all domains essentially regard selection decisions as:

*'It's all about judgement and that's judgement of the individual curator and judgement of the body of professionals in the museum' (P5)*

Senior managers are expected by respondents to have set the boundaries to collecting such as through policies and mission statements, but within their own areas of responsibility the respondents described being able to use their own professional judgement. Judgement is based on professional training, experience and specific knowledge and skills.

Whilst the practitioners were keen to emphasise the professional nature of selection they also recognised that it is essentially a subjective decision. Everyone is slightly different and exercising judgement is a personal activity:

*'It a value judgement and these sorts of judgements are horribly imprecise unfortunately ... it's very hard to say but that's just how it is [laughter] and I know that doesn't go down well and it isn't what people want to hear necessarily because it's probably closer*

*to the truth than some of the model based decision making that might be in the text books if you see what I mean.'*(P15)

Subjectivity is important during appraisal with the volume of digital material; where it is not possible to check every file, it is a subjective decision which sections to sample to check (P11). Then there is subjectivity in deciding what should be preserved; P7 gives the example of deciding between raw data sets or data sets after some form of processing. Further subjectivity is introduced when the creator is involved in the selection decision. There may be a tension between what the creators think is acceptable to archive and what the professional does, as they are judging things differently. Value and significance assessments are subjective even with clear criteria and procedures for assigning value:

*'So again value is a moveable feast depending on the time in which the assessment is made and the person who makes the assessment and the restraints on the decision that informs the value.'* (P21)

The suggested way through this is that there are clear criteria and procedures to make sure people are selecting consistently, although it is clear that there is a reticence in practitioners to refer to written guides and there is a clear reliance on professional expertise. This can be developed with training but it can improve with more experience, though P7 does point out that continuing professional development may not be easy if you are a lone practitioner with no-one to supervise you or to ask questions of. Whilst training is important, there may be problems when the practitioner conducting the appraisal is not experienced:

*'The problem with professional training [here referring to having completed a post graduate qualification] is you get told what to think and newly qualified archivists are often 'oh well I would keep this, this, this and this and bin the rest of it' because I haven't been told to keep that.'* (P7)

#### **5.1.4 Legal and ethical behaviour**

Being professional includes conforming to ethical behaviour and following legal requirements. It is partly defined through knowledge of and adherence to professional codes. The organisations which represent professionals in the cultural heritage domain– CILIP, ARA and MA for librarians, archivists and museum curators respectively – all have codes of conduct or professional ethics (section 2.3.5). The respondents also reflected a strong personal identification as an ethical professional:

*'I think it's making sure especially with professionals its making sure that that is at the core of what you do and it's got to be something that is almost kind of biological and in*

*your DNA I suppose....it's a question of being a good professional really and making sure you are doing things for the right reasons.'*(P12)

For archivists in particular being a professional means that you have been trained to be objective in appraising material. Being professional to the archivists meant being detached and setting aside personal preferences, reflecting the traditional view of the archivist's role as a neutral collector (Jenkinson 1937, p.149; Duranti 1994, p.343). The principle of objectivity leads one archivist (P7) to talk about her institution collecting records of the British National Party, regardless of what she feels about their policies. Ethical concerns of library respondents reflected similar concerns to those from other domains, including due diligence to ensure the provenance of items is reliable and that those selling items to the library are the legal owners. Library respondents also mentioned ethical problems of providing access to potentially sensitive or offensive material. The response of the library that owned a controversial book was to place a barrier in the way of casually looking it as the book was potentially inappropriate or offensive' if users wanted to see the book they had to ask for it to be removed from a cabinet (P17). None of them mentioned a similar issue with digital material though. If there is material that is potentially sensitive for a particular community, or if the library wanted to use it in a different way, then one library manager (P19) claimed she would try to include them in discussions about it to share understanding. Provenance of material was also cited as important by museum practitioners where they must ensure that the donor is the genuine legal owner of the material, or at least being able to show that 'due diligence' has been followed i.e. that all reasonable enquiries have been made to establish ownership.

When asked about ethical issues in selection for digital material, respondents also included compliance with the law relating to collecting, especially intellectual property rights. Legal and ethical issues are clearly closely associated to respondents. By complying with legal requirements the respondents felt they were acting in an ethical manner and no one questioned this or gave any examples of when it would be ethical to not comply with the law. The library respondents stressed compliance; being seen to comply with the law and working ethically enhances the reputation of their library. A library manager mentioned working in a transparent way for the same reason; if you want people to continue to donate material to the institution they have to have confidence in how you manage material. This library manager, who is highly engaged with digital material, saw no difference in the ethics involved with paper and digital, and this was reflected by the majority of other respondents. One particular factor that was highlighted as important when considering digital material was authenticity, which an archivist respondent linked with ethical behaviour; providing authentic original materials is seen as a show of honesty:

*'It's a question of sort of being... of objectivity and honesty to the users, staff, customers here. There is an ethic in providing the original record and not a substitute...'* (P9)

Archivists were clear that they would want the original item. Authenticity is important where digital material is used as evidence; it is easier to change digital material undetectably than with paper records (P7) and this may undermine its value as evidence if it is not demonstrably authentic. Here the archive respondents echoed the museum curators when they expressed concern about changing material, asking is it acceptable to change the format of material for preservation purposes or does this change the item too far from what the creator intended? The respondents suggested that methods of demonstrating authenticity may need to evolve; P10 considered it more likely that authenticity will be demonstrated by procedures and showing a 'chain of custody' as much as in supporting evidence and contextual information. This is similar to views found in the digital preservation literature where authenticity is demonstrated through using metadata to trace changes rather than insisting that the item remains unchanged (DPC 2008, p.24; Salza et al 2012, p.26).

### 5.1.5 Knowledge and skills

Respondents were asked about the skills and knowledge they felt were needed to select digital material. There was a strong feeling that by being a trained professional or being experienced means that you have most of the skills needed, reflecting the similar view of the experts (section 4.1.2). According to P13:

*'So what we say to anybody who is a trained archivist is that as a trained professional or experienced archivist you have eighty to ninety per cent of all the skills you will ever need to work with born digital.'*

This seems to exclude those with librarian or curatorial skills. However by examining the skills practitioners thought were necessary it becomes clear that many they mentioned were personal qualities, rather than skills that someone could learn through training. The qualities or abilities cited include:

- Initiative
- Persistence and curiosity
- Common sense (this was a popular response)
- Impartiality
- 'Political' awareness
- An open mind
- Ability to manage relationships
- Passion about your subject
- Enthusiasm

Trainable skills which the respondents mentioned include:

- 'Historical sense'
- Cataloguing or description
- IT literacy (not only for digital material)
- Strategy and financial management (for managers)

In addition respondents considered knowledge of the following to be necessary for practitioners, all of which are part of the current skill set for practitioners:

- Subject knowledge about your area of responsibility
- The current collection and of collections elsewhere
- User requirements
- The institution's purpose or mission; requirements of the business
- The collecting or collection development policy
- What you want to achieve
- Where to go for help and advice

Only one other skill was identified specifically relating to selecting digital material and that was the ability to keep up to date with developments. Also in the practitioner's responses there was a focus on technical skills, such as how to migrate material, and although this is not clearly linked to selection it seems that there is an assumption by some that these skills may be necessary.

Three different potential skills sets for digital material were suggested by practitioners. Firstly, having curatorial skills and working closely with those who have the more technical skills i.e. separate skill sets. In this view the need is not for extra skills (although a general awareness is useful) it is for appropriate workflows to be in place so that the practitioner knows what to do with digital material and who to pass it to for management. There is an underlying assumption here that the necessary skills for working with digital will not differ greatly from those for traditional material as it has to be selected and managed using the same underlying principles. Some respondents felt that there will be more specialists with technical IT skills and there would be a digital specific skills set for electronic records which centres on how to migrate material, recording metadata, deciding on format and so on. Amongst many respondents there seems to be an assumption that these specialists will be archivists who have chosen to specialise in digital material, rather than computer scientists working in archives. Alternatively special skills could be delivered through third party services. This reflects most closely the assumption by experts that a digital curator profession will develop. The third type is that practitioners will all need to increase their skills to include IT to manage digital material themselves; these may be termed hybrid skills sets. P12 says that the range of skills archivists will need will be wider:

*'so what I am going to need to do in the future is as well as help someone with palaeography in the search room I'm going to need to know how to convert a file into an open source format and need to know how to deal with that... metadata and stuff so the skills that I need are going to change because the material that I'm dealing with is changing'*

This is understandable as P12 works as the sole archivist in her institution. According to this view, the skill set for lone workers will have to change to encompass more skills, at least on a basic level as they may not have anyone to immediately ask for help from. Interestingly, a manager from a large national organisation claimed that a digital archivist will need a combination of traditional archival skills such as appraisal and digital specific skills such as how to use particular digital tools. Here they were describing a person with a mix of subject skills and knowledge, and IT skills. In their institutions it may be that staff have the opportunity to expand their skills because they are interested in doing so.

The respondents were also asked whether they felt their institution had the skills necessary to manage digital material already. There was a range of responses which reflected both the level of engagement that institutions have at the moment with digital and also which skills the respondents were referring. Those who answered with a definite 'yes' were those who were already highly engaged with digital material. There seemed to be confidence among these respondents that even though they did not have the skills themselves, they have people who they know they can go to for technical questions. These respondents tended to be from large organisations; P13, who works in a large well-resourced library managing digital assets, stated that he knows the right people to go to. A team as a whole may have the right skills, not necessarily all in the same person (P15). Technical and collection skills need not reside in the same person but managers do need the people with each skill set within an organisation to talk to each other (P19), reflecting the observation by Runardotter et al (2011, p.76) that responsibility should be shared between archivists, managers and IT personnel. It may be that the separate skill set model is more common than the others, perhaps because the digital curator role has not yet had time to develop. As seen in the literature review (section 2.3.3) there are few courses as yet focusing on digital preservation but if these increase then the role may develop in the future.

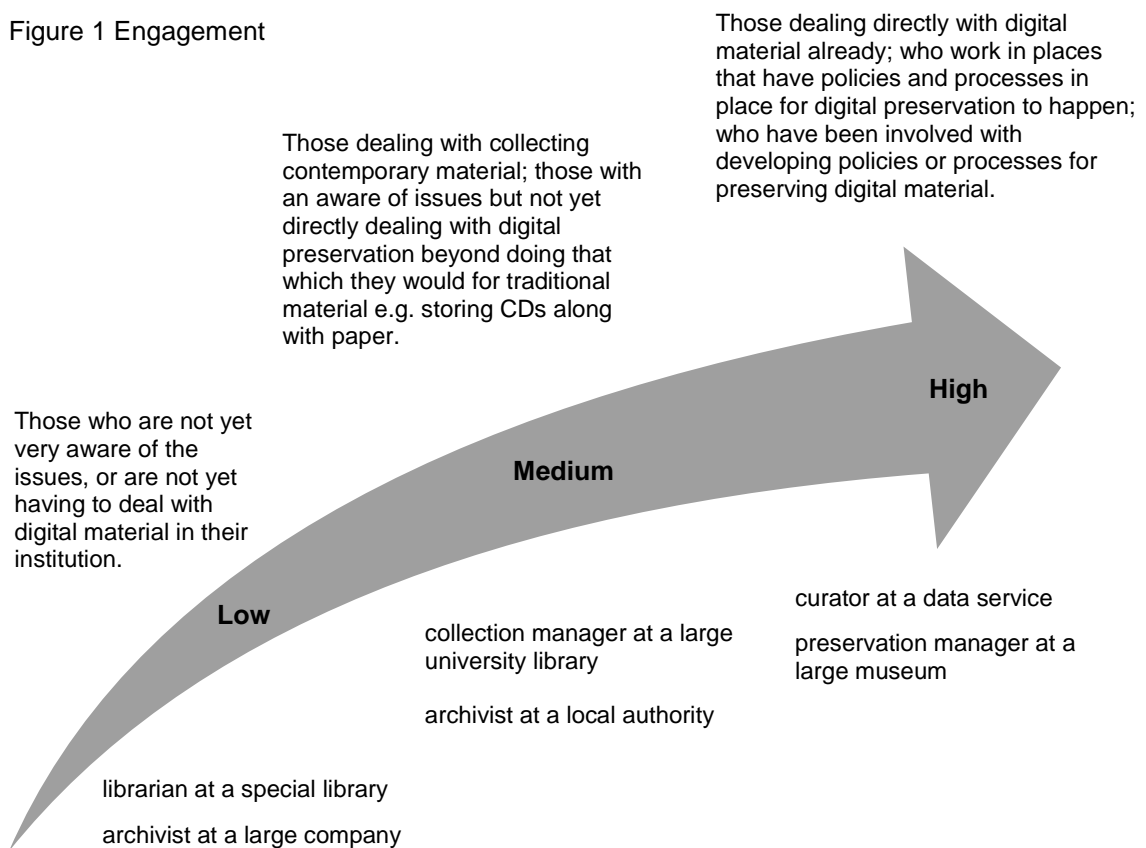
Amongst the managerial respondents, staffing for selection and preservation was of concern. P1 described problems arising in her institution because there is a difficulty in retaining staff skilled appropriately for digital preservation activities; due to financial constraints most people are not offered permanent contracts. She worried about not having enough staff to undertake all the necessary activities for digital preservation, as do other managerial respondents. More broadly, P14 identifies a reduction in his staffing budget as a limitation on the ability of his

archive service to proactively collect material, rather than waiting for material to be donated; here he is referring to both digital and traditional material. In addition people who are being recruited now may not have the right technical skills. According to P15, asking for people with digital preservation experience or qualifications does not always attract people with the right personal qualities, such as those listed in section 5.1.5. There is not yet a 'digital archivist' profession, though P10 suggested that the general training of archivists in universities would have to change to include a digital component. The training received by the respondents was more often 'on the job' rather than through any formal training courses or education.

### 5.1.6 Engagement

Throughout the interviews it became clear that some respondents were more knowledgeable and engaged with digital material than others and this affected their responses. Engagement here refers to their awareness of the issues surrounding preservation of digital material and whether they have had to be actively involved with the planning or practice of digital preservation within their institution. It is recognised that due to the nature of the research and the sampling strategy the ordinal scale used here to characterise practitioners is not in any way representative of anything more than the responses of the practitioners in this study. Three groups of respondents were apparent; those with high, medium and low engagement with digital and these are shown in figure 1:

Figure 1 Engagement



The group of respondents that are highly engaged include managers with an overview of preservation in their institution; they may be involved in formulating workflows and policies. There is a mix of settings; the manager from a national museum (P22) is responsible for digital preservation as a consequence of his responsibility for the technical aspects of film preservation and his role has expanded to include digital material. The high engagement group includes large or national institutions or data centres which have specific responsibility for digital material. Their job titles may include 'digital' or 'data'. Medium engagement practitioners can be characterised as having an interest or awareness of digital preservation but aspects of their institution, such as having sufficient management interest or the technical infrastructure, are missing. Responsibility for digital in the medium engagement group seems to have come about as a by-product of their existing roles; they have had to become interested in digital material because their institution has begun to receive or collect it. The low engagement group includes people who are not yet very aware of the issues, or are not yet having to deal with digital material in their institution. These include circulating librarians, a museums officer, a company archivist, and a collector of specifically non-digital material. The issue of engagement is particularly seen in consideration of training needs and issues of anxiety and confidence in selection; findings relating to these are in section 5.3.

## **5.2 Stakeholder roles, responsibilities and relationships**

It is clear from the literature review and the interviews with digital preservation experts that roles and responsibilities are an important issue, so the practitioner respondents were asked questions relating to this. Respondents were asked who they thought the stakeholders were in selection. Table 21 in appendix 5 identifies all the types of stakeholders that have roles in digital preservation, as described by the practitioners in different domains. Some respondents worried that there may be more stakeholders in the selection and preservation process, such as creators and IT staff; it would not be only the practitioner making decisions about the collections. To them this means that there will be further committees, prioritisation, and views to take into account and so on, which will slow and complicate the process. A university archivist (P20) expressed worried that in the future for digital material selection will involve '*so many other people than the archivist*' and will not be based only on the archivist's judgement. There is a frustration at what respondents perceive as a lack of co-operation from other stakeholders. The relationships that were reported in the interviews varied from very positive, where the practitioner was able to gain advice from colleagues for example, to very negative, such as with IT staff or managers who did not support the practitioner; these are explored further in the rest of this section.

There is a strong similarity between domains in the stakeholders they identified, allowing an overall set of core stakeholders in memory institutions to be detected. The most important for digital material according to practitioners are:

- The practitioners themselves
- Colleagues
- Managers
- Senior managers
- Institutions
- Users
- Creators and donors
- IT staff
- External funders
- Other institutions

However some differences were apparent and these are also shown in table 21 in appendix 5. Archives and museums mentioned relationships with community groups as being important as sources of material, reflecting their reliance on donations. Whilst libraries may get donations, this is not seen as a regular source so relationships for this purpose were not emphasised in libraries. The local community is seen as part of the libraries' user group however. Also noticeable is that national institutions and commercial companies are not seen as stakeholders in museums; this is unsurprising as the museum sector does not have a recognisable national institution with leadership qualities, unlike TNA for archives and the BL for libraries. There is an ethical stipulation in museums (MA 2008) to not trade in items.

Analysis of the interviews with the practitioners reveals a distinct set of responsibilities for practitioners, managers, senior managers and the institution itself. The majority of the responsibilities identified are generic and apply to selection for both digital and non-digital material. Many of the practitioners were clear that it was the role of their institutions to preserve material they had collected:

*'We are responsible for preserving all of them because we have chosen to collect them therefore we've also chosen to preserve them.'* (P13)

The assumption of permanent responsibility once an item is in the collection that is applicable to non-digital material continues with digital material. Certain aspects are different however, in particular the influence and role of other stakeholders. The respondents were clear that the responsibility for digital preservation could be shared with other stakeholders, including creators and donors, users (both internal and external), IT personnel, other organisations, and managers. Sharing responsibility is seen by some respondents as a positive approach to digital material that makes it easier to select and manage. In their view, as the volume of material to be collected and preserved increases so should the extent of collaborative working.

Through analysis of the interviews the roles that practitioners regarded other stakeholders as playing were discernible. These were:

- Sources: those who supply or hold the 'anti-collection' from which a selection to collect will be made. These could include software vendors, collectors, creators and donors.
- Colleagues: those people who give advice either formally or informally which helps the practitioner make selection decisions or set boundaries and limitations on selection. These include professional organisations, senior managers and immediate colleagues.
- Users: those who may use the material now or in the future; the audience for whom the selection is taking place who may or may not be known.
- Collaborators: these may be funders, commercial companies or other organisations with whom responsibility is shared for preservation so may have an influence on selection for particular projects

### 5.2.1 Sources

Sources refer to the people and places that material for the collection may come from. Table 14 indicates different types of sources mentioned by respondents from different domains that may have an effect on selection activity. These clearly reflect the traditional sources of material described in the literature (see section 2.1), though illustrate the broad range of stakeholders in sourcing material.

A form of proactive collecting was described which involves making individuals or groups aware of your service through outreach activities, in the hope that they will then deposit material with your service at a later date. P12 describes this as:

*'So it's kind of proactive, but it's proactive with your sort of hood up if you know what I mean because you're doing it in quite a subtle way'.*

Whilst this may have no immediate effect, the impact on the service may be great in the longer term. According to respondents donations may make the selection process more complex as:

- It might divorce the material from its context
- It increases the amount of material to be considered
- Means that deposit status needs to be negotiated and agreements drawn up

Whilst these issues are important for non-digital material, they may be more so for digital. As seen through the literature review, collecting contextual information for digital material is seen as particularly important (section 2.2.1), the volume is much greater (for example Feeney 1999, p.11; Harvey 2007, p.9) and clear agreements are needed (CCSDS 2004). This may have

implications for whether the practitioners solicit donations to the same extent, although local archivists will still need to fulfil their role in representing their community.

Table 14 Sources

Type	Sources	Example	Archives	Libraries	Museums
Donations	public donations	gifts, bequests, loans	x	x	x
	asking for donations	for an exhibition; for an institutional repository	x	x	x
Buying	buying online	eBay	x		x
	auctions or specialist dealers		x		x
	publishers			x	
Statutory requirement	legal deposit			x	
	parent organisation		x		
	other non-heritage organisations	welfare organisations	x		x
Referrals	local informers and enthusiasts		x		
	other heritage institutions	loans, transfers, exchanges	x	x	x

Local informants and enthusiasts increase the availability of material to choose from as they alert the archivist to local material that may be collectable or at risk. This relationship is important to archivists in particular, but evidence of it affecting museum collecting was given by P5. If the source is a 'rescue mission' then the archivist or person sent to collect the material may have picked up irrelevant material, so the archivist will appraise the material after acquisition. P7 describes this:

*'... the times that we do is when we have taken things in on the spot where an office is closing or a schools closing down and we've been told to come in and clear it all out the records where we might put it somewhere and then later think we'll have to go through that more thoroughly.'*

A librarian at a national library also mentioned participating in rescuing archival material, which then forms part of the library's special collections.

There may be a statutory requirement for an archive in particular to collect certain material, such as records of the local authority for local authority archives, although P22 explained that his museum is a place of deposit for material from a certain group. However there may be difficulty in receiving material from the parent organisation:

*P12 'but it's hard enough to get hold of paper records and to get hold of electronic records is even harder especially from your parent body...'*

Although they have to accept material so there is little selection by the practitioner prior to receiving it, the archive still appraises after receiving the material. According to traditional archival theory this should not need to happen (Jenkinson 1937) but according to respondents, creators do not always recognise their responsibilities for accurate record management so the deposits may have duplicate or irrelevant material attached. P22 has a similar problem, that the group from whom they receive deposits supplies them with too much material as there is nowhere else that takes responsibility for it. Because the depositor's material is now in digital form the museum has had to become engaged with selecting and preserving it.

External colleagues also act as sources of material. Many respondents gave examples of contact with external colleagues leading to a loan or acquisition of traditional material, though none mentioned this with digital material. This is unsurprising as digital material can, with the appropriate permissions, be accessed from anywhere so there is no need to loan material in the traditional manner. Material unsuitable for one institution may fit the mission or purpose of another and through close relationships practitioners are offered or exchange information about material. Practitioners also identified the role of commercial organisations in selection and preservation:

- As drivers or funders of activities, such as digitisation of particular treasured items which could then be turned into a book or other product
- By providing material for purchase
- As creators of software or other digital material to be preserved

Creators are an important source of material for those institutions which collect contemporary material and various outreach activities with creators were described by respondents. The respondents recognise the importance of accessing creators whilst they are still available in order to gather contextual information. It is important to build relationships with creators in order to source material. P11 explains that in her institution creators are encouraged to talk to the institution prior to creation to ensure this happens, although she has the advantage of working in an institution that has a specific area of responsibility with a well-defined designated community making it easier to build a relationship. Practitioners who accept digital donations

described making preservation easier by influencing creators before deposit to use 'preferred' formats (confirming the suggestion by E2 that this happens) and to supply appropriate documentation or contextual information for their material (P8). Essentially the creator is encouraged to go through the selection and retention process themselves before deposit in the repository and these respondents were very clear that awareness needs to be raised in creators about their responsibilities. P11:

*'Essentially we want people to think about the whole process not just copy everything onto a CD and dump it on us.'*

This recalls the issue of the archivist interfering in the creation of the archive. Pragmatically this is a method of managing file formats and volume, but it is unclear to some archive respondents how acceptable this is given traditional notions of the archivist role. But working closely with creators prior to deposit raises questions about the archive's traditional neutrality and becomes an ethical question. Whilst it is seen as an advantage to be able to access creators early in order to access contextual knowledge, P20 questioned the intervention of an archivist in the creation of an archive by being involved with the creator early on. She says:

*'...the last place I worked in we worked with one...one literary... I always had the slight impression that he discovered another notebook because we would purchase it... that's a very unkind thing to say but that sort of more or less creating the archive on demand.'*

The conceptualisation of archivist and record keeper's roles in the records continuum model, where there is a recognition that record and archive keeping is a shared responsibility, is more flexible; this model was developed partly in response to the introduction of electronic records (Upward 1996), and it seems to reflect practice as described here.

### **5.2.2 Guides and limiters**

Stakeholders are able to make selection by the practitioner easier through certain actions identified by respondents. This includes senior managers, who can ensure support is provided through the institution, and creators and software vendors can help by choosing appropriate formats and developing appropriate products. Relationships with colleagues are important to respondents; colleagues may be immediate or within the wider organisation and be users of the material curated by practitioners. Practitioners noted that selection decisions are not made in isolation but with regard to 'guides'. Guides come in many forms – professional knowledge, technical standards, ethical codes, the current collection and legal requirements were all identified by respondents - but a common guide mentioned by many respondents were other people, especially colleagues, who may help to determine the value of material. According to respondents, the main role of colleagues is to give informal support to decision making by the practitioner. For example:

*'Criteria are not always written down but are in the head of the librarian – by talking it over with colleagues you can get extra guidance....' (P17)*

*'Some items may be more 'tricky' to make a decision about and so colleagues give extra help.' (P7)*

Respondents felt it is important to consult colleagues both as sources of information and as potential users. This can happen on an informal level where the practitioner can have a 'chat' about things with colleagues who have 'unofficial' knowledge, for example about a publisher or about other people and how they work (P17). For those practitioners that work alone, or as the only practitioner of that sort in an organisation, wider professional networks are valued as a source of support and advice. Many practitioners feel they are able to access support and technical expertise when they need it and are able to learn from others working in the field:

*'...there's a good kind of network in the [redacted] area, we will share policies and people will kind of give advice to other professionals and professionals rely on each other for advice, so you feel that you're not really working in a vacuum and you're doing something that fits in with professional work elsewhere.'(P12)*

Informal networks may be formed across institutional boundaries between practitioners who wish to pool expertise and gain support for their decision making. Co-operation is seen between institutions (Angevaere et al 2012, p.95; Portico 2011, p.15) but here the emphasis is on the value of individuals being able to collaborate informally.

Respondents recognised that limitations are set on collecting through policies devised or approved by senior management colleagues. They also recognised that the role of senior managers includes setting boundaries to collecting by formulating high level criteria, institutional aims, and the mission of the institution. They also limit selection by controlling and authorise the allocation and prioritisation of funds; this role is the same for digital and non-digital material and was seen by respondents as unproblematic in the institutional context. A further important role for senior managers to provide a mandate for preservation; this is necessary to drive engagement with digital material forward within the institution, which helps practitioners to source material and to emphasise the importance of their role. P16 described how senior managers supported the development of an institutional repository in her organisation:

*'When this was all developed we had to get a mandate from the senior leadership team because otherwise you know no-one in [...] would ever have done anything with it. So they had to come out and support it and say we want [...] to do this, so they had to do it.'*

However it became clear from further analysis of the interviews with practitioners that the quality of management support may affect selection. An example of the difficulties that can arise through the organisational structure was described by P4, who is a company archivist. She links the sometimes difficult relationship she has with the IT department in her organisation to the organisational structure of the company. P4 states she has no senior management 'champion' and therefore she has no influence over how the organisation selects and manages its electronic records. The archive is part of the marketing department and digital records are the responsibility of IT. She refers to this arrangement of responsibilities as a '*cultural misfortune*' and a '*historic accident*'. The core work of IT involves financial records and commercial security but they also perform what they consider to be 'preservation' activities, such as backing up servers, and feel they are 'owners' of all digitally created material. Thus they had recently 'archived' all electronic files that had not been amended for a year, without consulting the archivist, including electronic material she had collected as part of her duties in documenting the company activities. In her words:

*'All my research files, scanned images, reports etc disappeared from the system and it took me days to persuade them to restore the material and to promise never to do that to my folders again.'* (P4)

Because the IT department in P4's company has a great deal of autonomy and its management is not well integrated into the management of the company as a whole, decisions are made independently and without consultation. Some organisations have senior managers who act as 'champions' of the institution. Others do not have that representation at higher levels and so it more difficult to get the needs of the institution addressed. P12 works in a local authority that does not have a records manager. This makes it more difficult for her to obtain paper records and has increased her anxiety about whether she will be able to obtain digital records.

The need to work more closely with IT colleagues is also perceived by respondents as being potentially problematic. As noted by Oliver, Chawner and Lui (2011, p.314) this has not been greatly examined in the literature. Whilst IT staff may not have responsibility for selection, they may have responsibility for preservation and therefore have influence over what is kept. In one commercial organisation the archivist explained that '*IT sees itself as owner of all digitally-created material, and will not be budged*' (P4), so she had little influence over the digital records. Many respondents were noticeably keen to differentiate themselves from technical people; P5 clearly stated '*we are historians, we're not technical people*'. IT personnel are seen by respondents as having very different views of the material and of appropriate processes and procedures, which can lead to difficulties. Some animosity was revealed in the practitioner interviews, which implies that working together needs careful management. P19 described how in her institution the archivist's workflow for paper had been translated into a digital preservation

workflow. Collection staff were concerned with making sure of *'the purity of the collection at all cost'* i.e. keeping records together and in context, so preserving the integrity of the archive. The IT people did not understand why the procedures had to be so complex, as they did not understand the agenda of the collection staff. The problem here seems to be one of a lack of a shared understanding of the function of an archival collection. This example illustrates well the conclusion from Oliver, Chawner and Lui (2011, p.321) that the cultural differences between IT people and record managers are very wide. In order to overcome difficulties it is necessary to engage both IT staff and curatorial staff in dialogue to increase understanding.

### 5.2.3 Users

As pointed out by E2, 'users' are not necessarily current users, they may be future users. The institution may not have either a clearly defined community nor be near their users, especially in an online environment where they may be geographically distant. Respondents emphasised that they take account of the external stakeholder when selecting, but in response to follow up questions none were clear about how they went about this. Gathering user views seems often to be carried out informally, unsystematically or by chance and there is an assumption by respondents that they know already what users might want or how they may use the material. They clearly demonstrate an expectation of the continuation of their traditional role as a 'proxy' for users in selection.

Groups of users mentioned by respondents include:

- the public, including families, schoolchildren and tourists
- researchers e.g. academics, family history enthusiasts, students
- internal users e.g. colleagues, learning teams, exhibitions staff
- other organisations that use heritage resources, such as industry specific companies or heritage organisations
- other library, museum or archive professionals or institutions
- people in the future

These are broad categories and it highlights the potential difficulty with identifying a discrete 'designated community' in many institutions, as implied by E2 (section 4.1).

Respondents from all domains mentioned collecting or preserving material for students or researchers, now and in the future, and that different types of researchers will have different requirements. From the description of the influence of users on selection it is clear that users affect the value of material, as changes to user demand increases or decreases the value of the objects or collections. In order to satisfy the potential demands of researchers, it seems that much material is kept just in case it is needed by them (P17; P2). Traditional archival values

(Schellenberg 1956) include consideration of possible value for future users, but this strategy is also apparent in the collecting described by respondents in other domains. However P19 explained that in her experience libraries are unable to collect 'just in case' a user may want to access digital material any longer as there are not enough resources, so the institution must focus on user demand to lead decision making. Many of the practitioners insisted that collecting is led by user demand. In their view users are able to make suggestions about what to select and local informants or enthusiasts may alert practitioners to potential new material. Some respondents demonstrated how they gathered information about their users through audience breakdown statistics, surveys, focus groups, or by requesting feedback both directly and from activities supported by the institution. The respondents felt that it is the existence and quality of relationships with different user groups that enables them to meet user needs, both by understanding their needs more fully and in anticipating future needs. It is also clear from the interviews that those institutions with a smaller, more clearly delineated user group, such as a particular academic group or a particular company, are able to build closer relationships their users and tailor selection and preservation activities more closely to their needs.

#### **5.2.4 Collaborators**

When the respondents spoke about relationships between institutions, they referred to collaborations, whether on a short term project basis or for the longer term. Working together in collaborations is important because it allows individuals and institutions to share responsibility, costs, skills and best practice (Lunghi et al 2012, p.214). In the view of those respondents who had been part of formal collaborations, collaborations require:

- a shared responsibility between collaborators
- 'give and take' between collaborators
- trust between members
- risk management; for example a library manager explained how her institution has kept their own infrastructure to minimise risk if the collaborators infrastructure has problems.

Respondents expected that because of the resources needed for digital preservation collaborative working may increase, especially as institutions develop specific expertise in preserving digital material, so collaborations between different bodies may produce forms of outsourced support; networks or centres of expertise might be established to deal with digital material. Whilst most respondents who mentioned this view of future centres of expertise were positive about the potential for collaborations, P20 expressed reservations. She asked if archives would lose their attachment to place and more obscure but locally important material may not be collected. It seems there may be confusion here between 'area of responsibility' and 'area' of physical space and storage. Digital material is not limited to the physical space of an

archive in the same way as paper; it may be stored on non-local servers, but this does not mean that it is not owned by a specific institution. Distributed management might be arranged, controlled through the use of policies, standards and contractual agreements, and items stored at a distance could be included in the collection in a virtual way, though as E2 noted the issue of trust in third party services is not yet completely assured. It also reveals a concern that collaboration in infrastructure and expertise may lead to collaborative selection, where the interests of smaller or local institutions are not met. P20 is an archivist and so may not be familiar with collaborative collecting initiatives; collaborative collection management has long been useful in libraries (Johnson 2009, p.265; Day, Pennock and Allinson 2007, p.4). Collaborative collection management initiatives do not necessarily lead to collaborative selection, for example, in using the LOCKSS service libraries create their own archived collections.

### 5.2.5 Stakeholders and roles

Further analysis of the interviews with practitioners reveals the roles that each stakeholder group is seen as potentially having, which are shown in table 15:

Table 15 Stakeholders and roles

	Guides	Limiters	Sources	Users	Collaborators
Colleagues	x		x	x	x
Managers	x	x	x		
Senior managers		x			
Institutions		x	x	x	
IT staff		x	x		x
Creators and donors		x	x	x	
Other institutions		x	x	x	x
External funders		x	x	x	x
Users			x	x	

There are some notable aspects to this table. There are limited roles assigned to users; the practitioners did not view them as collaborators in selection. There are few guides; advice is mostly taken from immediate managers or from colleagues either within the institution or from external bodies such as professional organisations or those who have particular expertise. This seems to reflect the sources of advice that are used with non-digital material. IT staff are not viewed as giving guidance, perhaps because in most institutions they are not well-integrated

into the selection and preservation process. IT are seen as providing limits to selection through absence; if they are not working with the practitioners then the practitioners are not able to source material in some cases (P14), be confident that what is selected will be safe (P4) and by not supporting or providing the appropriate infrastructure and tools. Difficult or non-existent relationships with IT cause anxiety and limits working together, further hindering selection. Of the limiters senior managers, through setting policy, and external funders, through conditions of the funding and influence on the priority given to different actions and material, are viewed as the most influential by practitioners.

### 5.3 Reactions to digital

It became clear throughout the interviews and from further analysis that the respondents had different emotional reactions to digital material. The most positive response to digital came from those respondents who see managing it as a continuation of their existing skills and responsibilities. This is summed up in the attitude of P13, who works in an institution that is very engaged with preserving digital material:

*'So just because its digital it doesn't mean you don't have to do all that donor creator contact stuff, all that wooing of the donor creator, we still have to do appraisal, we still have to do arrangement, we still have to do description, it still has to fit the collection policy. It's just stuff.'*

Here digital is not seen as something different; it is *'just stuff'*; P13 has reached the acceptance of digital as stuff that E8 hoped practitioners would. The engaged practitioner views the processes and procedures needed for digital material as having a connection or a continuation with those they are familiar with for analogue material. This applied to the skills needed, the criteria that are used for selection and the underlying principles behind actions. Archivists in particular were keen to stress that fundamental archival principles relating to appraisal would not change with the change in media. Differences were acknowledged in the technical skills needed and the type of storage needed, but they did not regard these differences as insurmountable. Professional judgment still has to be used to select and appraise material. P2 was positive about preserving digital material if *'a robust kind of focussed approach'* is taken to working out best practice. P11 was also positive but felt greater consistency in selection and deaccessioning is needed in her institution through the use of strict criteria. Some respondents recognise that what they can do with digital at the moment is only *'their best'* (P12). This is a pragmatic approach and P21 in particular is reconciled with this:

*'But at the end of the day we have limited resources and we have limited time, limited everything so we just try and do the best we can. I like to think we are just trying to be as honest as we can about what we can do...'*

P19 also recognises that what she is trying to do is balance resources against the desire to produce a '*perfect archive*' and this is where there is a need for '*rational criteria*'. The importance of clear criteria to guide decision making was often reiterated.

The interviews indicated that digital material produced an anxious response in many of the respondents. Few were not anxious at all and these tended to be those whose institutions were focused on digital, such as data centres, or those whose institutions had not yet engaged with the problem of digital. For the majority of respondents, the introduction of digital material produced anxiety:

*'once I get something new once I get that first memory stick that is the whole of somebody's collection in an electronic form then I'll kind of panic I think.'* (P12)

Once the material has been accepted by the institution, concern turns to controlling the care of material. The relationship between practitioners and IT staff is relevant here. A few respondents worried that the IT department could choose to dispose of digital material without consulting the curator. When discussing digital material and digital preservation, P2 describes it as '*a giant hulking problem [that] is kind of eclipsing the horizon*'. P12 confided that:

*'... I could blag it a bit but I wouldn't feel as confident as I do dealing with traditional material.'*

These responses reflect the anxiety and negative feelings that digital material may cause. Viewing digital as a continuation of non-digital material, with similar issues, seems linked to lower anxiety. Those who feel that digital is 'different', something other, expressed greater anxiety. Not knowing what to do with digital material engenders anxiety and having control of the acquisition of material was of concern to some respondents.

It seems that in many institutions as yet there are not the workflows or deposit arrangements worked out to deal with this, or with unexpected deposits of digital material, causing anxiety. This needs to be planned for and processes put in place so the material can be dealt with appropriately when it is donated. This not only applies to unexpected donations, but also to material in unexpected formats. P10 describes finding '*strange disk cartridges*' in an archive that was donated, which leaves him with the problem of finding someone with the right equipment to read it. But it was heartening to also be told that there is a sense of community, an awareness and comfort in knowing others are struggling with similar issues; both P5 and P10 made it clear that they do not feel alone in dealing with preservation issues; both are in touch with practitioners and experts in other domains.

When comparing the factors that cause anxiety between curators and managers it becomes clear that curators are concerned with having selection or appraisal procedures in place to deal with material, whilst managers are worried by a lack of understanding of, and confidence in, the technology. It is also possible to determine issues that create anxiety between those with high, medium and low engagement with digital preservation. High engagement practitioners are mainly concerned with technological issues, including keeping up to date with developments in digital preservation:

*'you know when you're trying to actually achieve things in terms of preserving the data it's quite hard to find the time to sit down and work through all these tools and things'*(P11)

Anxiety is caused by not fully understanding the technology or what is coming next, along with having confidence in the technology to preserve in the very long term. Being able to manage large scale digital objects in the near future when these need preserving and whether the resources will be available to migrate current digital material when necessary were also mentioned as worries:

*'...the difficulty will come in five or ten years when the first generation of stuff has reached the end, format obsolescence when we need to migrate it and the money may not be there.'*(P22)

The high engagement practitioners were also worried by the effect of current legislation and business models on long term preservation, for instance knowing how to comply with data protection or copyright legislation when dealing with new forms of material, such as emails.

A concern of the medium engagement group is having appropriate guidance and procedures in place. Not feeling adequately prepared for the expected influx of digital material creates anxiety. This group recognises what it is they do not yet know and feel the need for guidance to help them when they begin to deal with digital material in volume or on a regular basis. Guidance on the following was mentioned by respondents as being particularly useful:

- Control of sensitive information – with paper it is easier to isolate sensitive material or put barriers in the way to prevent inappropriate users gaining access, but this is more difficult with digital material
- Manage access or preservation of material on different or older formats or media
- Manage IPR issues, such as agreements with owners and donation policies
- Determining which versions of something to keep
- Managing the expected volume of digital material, especially by having appropriate appraisal procedures and policies in place

Guidance needs to be created on these issues in order to alleviate anxiety. In order to prepare for managing digital material effectively the right skills and institutional infrastructure are necessary:

*'If somebody brings me a hard drive what am I going to do with it? I need to be equipped to be able to deal with that.'* (P12)

Additionally the respondents with a medium level of engagement identified not having a clear person with lead responsibility or a clear line of management responsibility for records and archives within the parent organisation as being a source of anxiety. For archivists there was also the worry about how the archive can ensure it receives electronic records in order to preserve them, especially when many reported difficulties in receiving the appropriate paper records.

The lower engagement group also demonstrate a certain level of anxiety but not yet about digital material. Not unsurprisingly their concerns centre on wider issues of keeping their service funded by demonstrating relevance to their users and management. A review of the service due to the introduction of electronic material was cited by one worried librarian as it may potentially leading to a loss of staff or resources:

*'They've cut the library staff drastically because it's going to be electronic and apparently that doesn't need any managing [laughter] we can already prove to them that that's not...'* (P16)

### **5.3.1 Technical challenges**

According to respondents, volume is seen as one of the biggest challenges of digital material; it is much easier to create multiple copies and versions of something. P15 likened the change in the scale of processing needed as from a cottage industry to an industrial system. They anticipate that it will take a long time and a large amount of resources to not only to select which material to keep, but also to manage it by cataloguing it to make it useable, and to check for data protection requirements. P21:

*'As for what happens to all the hard drives and material that sits on people's shelves with images in, in curators' offices, there's clearly going to be important, useful and significant stuff in them but how you manage that without looking at it i.e. devoting thousands and thousands of people hours to it, I don't know...'*

Archivists in particular were concerned with appraising digital material, saying there is greater difficulty with digital in appraising at a detailed level. Some archive respondents made clear that more digital material is kept because it is harder to appraise and more resource intensive to sort the good material from the bad; it is also more difficult to decide which version of something to keep. It is more difficult to appraise or weed out duplicates on an individual file level as there is not the time or resources to appraise each file. P7 explained that it is much easier to keep a whole CD than look through all the files on it to see which ones are useful. More 'rubbish' is being created as it is so easy to do, but within that rubbish there might be important information. So for P7, the best option is to keep most material at the moment for a certain amount of time 'just in case'. According to archival respondents, whilst basic appraisal principles remain they need to be examined for relevance to electronic materials and procedures may have to evolve in order to deal with the greater volume. Suggestions given for managing selection from greater volumes included using stricter criteria for collecting and deaccessioning and more sampling of material will have to be undertaken. Collecting policies, understanding workflows and costings will become more important to allow greater control over the material (P19). Respondents felt that the processes of preservation will need to change to take account of the higher volume of material and the greater time and costs that will be needed to manage that volume. What is clear from the discussion of volume is that it is recognised as an issue by archivists in particular, reflecting the development of different forms of appraisal in reaction to increasing volumes of non-digital material, such as functional or macro methods of appraisal (Cook 1997, p.47).

Format is seen by many respondents as a difficult issue for several reasons. Firstly the variety of formats has increased so the range of items the practitioner must be able to manage has increased. This increases the resources needed to manage it; when asked what they thought would be influences on selection or preservation in the future, P21 mentioned:

*'The format of the records – whether the information can be captured and maintained, whether costs will be prohibitive'*

It should be pointed out that many respondents used the term 'format' to refer to media also, which may be unhelpful for them.

### **5.3.2 Significant properties**

The respondents were asked about significant properties of digital material which they considered important when selecting for preservation. This question proved a difficult one for many of the respondents as the term was unfamiliar to them. The researcher provided a brief explanation of the term where necessary. Even so, this was not a successful question to ask most respondents. Those who are highly engaged with digital material were more likely to be familiar with the concept.

The responses from archivists concentrated most on the content of items, the information that they contain. P14 draws the distinction between museums and archives as:

*'I think it's probably one of the ways in which we differ from museums we are information providers, we're not interested in collecting examples of things, it's what the documents contain.'*

The text, structure and formatting of the document as well as the content were mentioned as important. P11 (a digital curator) explained that in her work for significant properties there is a 'scale' of importance, with the content being the most important and further down the scale are qualities such as page numbers. If they are not exactly the same when the item is migrated this may be an acceptable compromise if it is too difficult to put right. This seems a pragmatic way of using a complex and ill-defined concept to assess the outcomes of migration. In contrast, P15, who is highly engaged with digital material and digital preservation, is more concerned with practical issues and solutions in order to do the day to day task of preserving digital material, rather than with what he calls 'academic' concepts. He viewed the concept of significant properties as found in the literature with some disdain, declaring that it *'...is of almost no practical value at all...'*

The library respondents took the view that the carrier was of secondary importance to the information it contained. Their focus was very much on preserving the ability to re-use the material, though not necessarily by using the original software; being able to provide continuous access to the material was mentioned by all the library respondents:

*'I mean the significant properties, I know the look and feel aren't top of our list, it's how you provide access to the actual information that's there...'* (P8)

For museum respondents the concept of significant properties had two aspects. Firstly the museum would want to collect and preserve an item in an aesthetically good condition. Secondly the condition of the artefact may in itself tell us things about its use or history. P5's museum role focuses on computing, so she is concerned with the hardware rather than preserving the data:

*'We're not so interested in the information that's stored on the floppy disk...the information on the disk is the icing on the cake. We'd love to have it, it's nice and it helps, but the cake itself is the object and it's how it looks, how it feels.'*

P5 also makes the point that the significant property that is most important depends on the needs of the user at that point. For the casual visitor if a video game is running a little too fast

when displayed it may not matter but to a researcher looking at the experience of playing the game it might.

The response of the practitioner to the question of significant properties, the lack of interest in the question and the lack of understanding of the concept implies that for practitioners 'significant properties' is not conceptualised in the same way as in the literature (see section 2.2.2). The concept as described in the digital preservation literature is of little use to them in their day to day work. This may be because they have not thought of material in this way before, or because someone else in their organisation takes responsibility for considering significant properties. It was clear that even when the concept was briefly explained it was not seen as having an impact on the work of many respondents.

#### **5.4 The selection process**

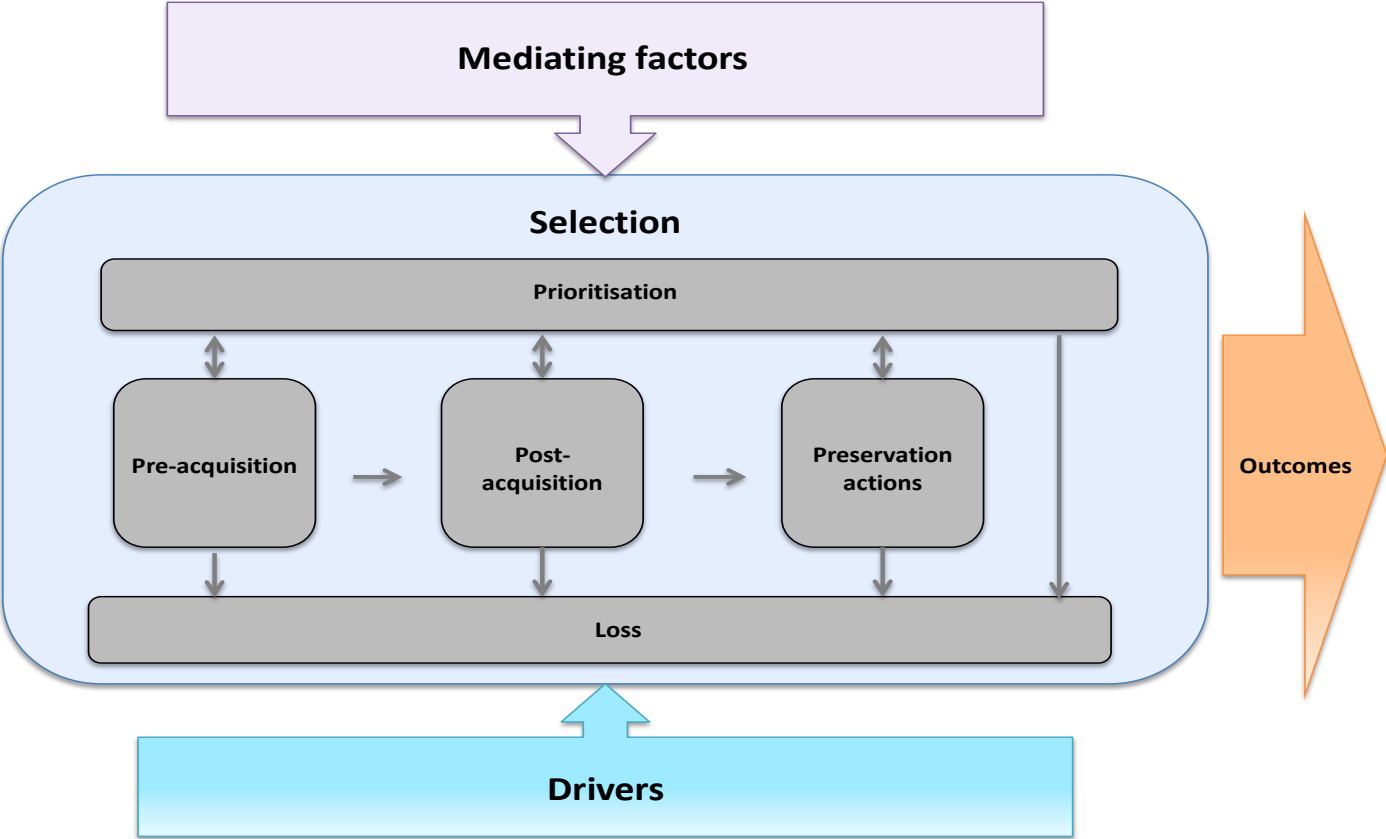
The literature discusses a core set of processes which constitute selection for both traditional and digital material (for example Johnson 2009, p.109; Winsor 2009; TNA 2013). The respondents confirm these processes, describing similar activities consisting of:

- Becoming aware of the need for selecting, whether this is for collecting or for further activities following acquisition, and becoming aware of the material available to select from
- Deciding value –evaluating a set of potential material for inclusion into the collection
- Acquisition – acquiring the material for the institution, or the ability to use the material.

From the analysis of the interviews it is clear selection decisions are made about material at different points, not only prior to acquisition, similarly to the process with non-digital material (Feather 2006, p.13). This is illustrated in figure 2 on p.140 which represents an overview of the selection process for all domains. Collecting, acquiring or accessioning material into the institution or the collection includes pre-acquisition and post-acquisition decision making. The former may happen when a potential donor brings material to the institution to donate and is also the point at which much library selection happens, as the librarian chooses from which source to acquire material. Post-acquisition selection refers to decisions made following a process of appraisal or selection to keep some material that the institution has become aware of, had donated, or might want to use. Processes here include formal accessioning into museum collections and appraisal in archives after acquiring the material. Prioritising material for potential further actions also includes selection decisions; as there are limited resources in institutions items need to be prioritised for preservation or conservation actions. Further selection is described by the respondents as happening where items are chosen for specific preservation actions, either directly from the collection, such as where an item has been damaged and needs urgent conservation work, or from the process of prioritisation, such as

from a collection assessment exercise. The final point of selection described by practitioners is for deaccessioning, deletion or weeding collections where there is decision to lose the material from the collection as the material is deemed no longer of value to the institution and is disposed of in the appropriate manner. If material is not chosen at previous stages this can lead to loss. All of these selection points are prompted by drivers, affected by mediating factors and lead to outcomes.

Figure 2 Selection points



### 5.4.1 Drivers to selection

The experts and practitioners identified institutional and material drivers to selection. Table 16 presents drivers mentioned by respondents from different domains, highlighted in blue with A - archives; M - museums; and L - libraries. The final two columns presents the drivers as they were related to selecting material for inclusion in their collections, which may itself have preservation implications, and for specific preservation actions.

Table 16 Drivers

drivers	A	M	L	collecting	preservation actions
active use e.g. e-learning, exhibitions	x	x		x	
availability of resources	x		x		x
commercial/marketing potential	x		x		x
external funders and organisations	x	x	x	x	
imminent risk to material	x	x	x	x	
importance of the material			x		x
institution of 'last resort'	x		x	x	
legal or statutory requirements	x			x	x
mission or policy of institution	x	x	x	x	
opportunity	x			x	
protecting reputation	x	x	x		x
uniqueness		x			x
user needs	x		x	x	x

Many of the drivers are as expected and reflect similar drivers to those for non-digital material; availability of resources, legal or statutory requirements, and use for example. A driver mentioned by respondents from all domains focuses on protecting the reputation of the institution, as suggested by E3. P1 described a driver to choosing to preserve certain materials as '*avoiding the serious reputational damage that would be caused if this material became unavailable*'. Also expected are drivers relating to value; uniqueness and importance are closely linked as material may be important because it is unique. A curator talked about uniqueness being a driver for digital material not only currently but also in the future, '*when we have a piece of software and we've got the only copy and it's probably really important*' (P5). This is also related to the imminent risk driver where material is selected in order to protect it from loss.

Selection is clearly driven through relationships with outside influences such as users, other institutions, commercial companies or creators and donors. There was a frank admission from

P8 that there was a strong influence from external funders driving selection activities in her institution. Material in projects that are externally funded or that have an external organisation involved take precedence so backlogs in managing material, such as creating metadata, builds up; there is a danger that material that is not expressly prioritised will not get looked at and therefore be left unsorted and uncatalogued, and hence eventually unusable. P8:

*'Whereas things that you haven't made an explicit commitment aren't going to get the same sort of priority are they? There's a danger of course that they never do, but let's not go there'*

External funders as a driver is linked to 'availability of resources' and 'commercial/marketing potential' drivers as these may arise due to the influence of external partners in supplying additional resources or a new way of using material already held. P19:

*'Yes revenue earning is a big part of what we're going to be looking at ...we would use that money to supplement our own digital content. Because sometimes we get requests from publishing companies that want to digitise our collection so in effect they are paying us to digitise it so they can sell...'*

The 'opportunity' highlights the ad hoc nature of much collecting. The respondents from archives in particular often saw selection as something that they did when they had the time. 'Opportunity' relies on either serendipitous discovery of potentially useful material by the practitioner or finding out about material through local contacts or other organisations. This in turn depends on contacts with local organisations and individuals which can act as informants. This could be problematic for selection of digital material as by the time they have the opportunity they may not have access to the creator or other sources of contextual information. P8 describes an example in her library:

*'Well what was easy was that we sent an IT person down with a memory ability to get the stuff it took about 7 hours to copy everything across. It wasn't particularly technically difficult but it meant that we got everything even stuff that we had no appraisal we just took everything ... But we haven't been able to get access to it yet to get started appraising it or to start cataloguing it and the more time goes by, [...] has come to an end and nobody in [...] we can ask. We've only got minimal metadata....'*

Some respondents, especially from archives though also from a national institution, felt that they are 'institutions of last resort'. This acts as a driver as they feel they have less choice in whether to accept material. P8:

*'The main driver would be the importance of the ... material and in this regard the likelihood of this material not being collected or safeguarded by any other institution.'*

#### 5.4.2 Criteria

Respondents were asked about the criteria they use. These criteria form a 'core' set for collecting, although it was clear that because there were few museums represented in the interviews few criteria used in them were discernible specifically for digital material. The core group centre on the following factors:

- ability to preserve material
- adding to the collection
- duplicates of other material
- meeting the collection policy
- others already collecting it
- 'just in case'
- use
- copyright
- resources and costs
- value, including historic importance and uniqueness

These have similarities to those identified in the literature (section 2.3.1, table 1). However each domain focused more heavily on certain aspects. As would be expected, each focuses on fulfilling the particular purpose of the institution – museums use criteria to select items that are representative, tell a particular story, have an interesting design for example, whereas libraries use criteria that identifies material that they are able to catalogue, make available and is up to date for example. Whilst many criteria were the same as for traditional material, the following were given greater emphasis or higher priority by the respondents when they discussed selecting digital material:

- The ability of the institution to continue to provide access
- The ability of the institution to read, store, and preserve the material
- Format or medium
- Costs to preserve
- Rights ownership
- Duplicated material

As expected from the literature review (section 2.3.1), practical criteria are of greater importance when selecting digital material. These were also mentioned by highly engaged respondents,

who although they viewed digital material as *'just stuff'* (P13), used criteria relating to the material properties of the item.

Professional expertise means that practitioners are trained to use their initiative and not necessarily have written rules for selection. When discussing criteria used in selection respondents from all domains described having informal (i.e. unstated) criteria which they carried in their minds which were not written down anywhere; whilst most could point towards a collecting policy of some kind the respondents mentioned that they do not always consciously refer to it. The use of judgement is viewed a kind of professional instinct:

*'now I won't say that every decision we make is instinctive but at the same time it's very clear that when you put a bunch of archivists and people who have been around this place for a little while and have them make a decision they will look at the situation and say well this is obviously the right decision.... It's slightly disingenuous to suggest there's been a rigid execution of a decision tree in coming to that conclusion. All of our decisions can be justified by reference to decisions and policies that we have but the calculation is not mechanical.'* (P15)

Analysis of the practitioner interviews identified further criteria that are used to justify non-acceptance which are negative versions of positive criteria. None of the respondents stated that they had explicit criteria for what they would not accept, but analysis of the interviews revealed criteria that are used to deny selecting digital material:

- not in an acceptable format
- not having adequate documentation
- a potential mistake by the donor (archivist only)
- duplicates
- institution not owning rights (though this does not seem to be absolute if the risk is acceptable)
- donor not owning rights to all elements of the donation
- non-functioning

These further reflect the criteria suggested in the literature, in particular having adequate documentation or contextual information. Respondents recognised the importance in obtaining this for digital material although the absence of it was not an absolute criterion if the material was sourced through a 'rescue mission', such as described by P8 in section 5.4.1.

The condition of the material is used as a criterion for de-accessioning. If the item is no longer functional then there is little point in keeping it in the collection. P2 has some digital art items that he was considering deaccessioning *'just because they don't work anymore and they're un-*

*operational*'. This then raises questions of ownership – if the institution owns the items then they are able to dispose of the item, whether it is digital or analogue, if they wish. However if they are not the owners then permission would have to be sought.

The respondents were all insistent that they did not select on format, but format was clearly a criterion for non-acceptance (P8; P11; P13). Various terms are used by respondents '*preferred formats*' (P8) or '*commonly used formats*' (P13), all of which refer to controlling the formats they are prepared to accept. Respondents also claim that they do not select on the basis of format. However the experts identified format as an often unstated criteria (section 4.3.2), and it is clear that this used as a criterion by the practitioners. The respondents that accept digital material encourage their donors to use preferred formats. Because archives want particular formats that they can deal with more easily it is easier and cheaper to influence the creator before deposit rather than take a format that the archive finds difficult to manage, which then may lead to the problem of the archive interfering in the creation of the archive.

### **5.4.3 Outcomes**

The respondents were asked why they performed selection and preservation and a variety of responses were received. Analysis of the interviews revealed further outcomes from selecting material for preservation. Outcomes link strongly to the influential factors identified throughout this chapter, demonstrating the effects that these factors may have; they also link to the preservation goals that an institution may have and indicate positive implications of selecting digital material. Many of these are similar to outcomes from preserving non-digital material.

Firstly there is a professional outcome; by selecting material practitioners are able to fulfil their professional and ethical needs. As identified in section 5.1, practitioners are emphatic that selection is a professional responsibility and that it can only be done by someone with the appropriate training and skills. They clearly derive satisfaction from performing their roles professionally so for practitioners a positive outcome is when this happens. Secondly there are institutional outcomes, including mitigating risk; protecting reputation and investment; fulfilling legal obligations and the mission of the organisation; fulfilling agreements and contractual obligations; and continuing the institution. The latter outcome was of interest to library practitioners in particular. P13 explained that in his opinion, selection helps to demonstrate the relevance of the institution, which not only justifies its current existence but also into the future. If the library stops collecting material then it will cease to be relevant so as more material is produced digitally then the library has no choice but to collect it in order to continue to exist. Similarly P16 described selecting electronic material for inclusion in her library through a particular project as a way of demonstrating continued relevance to senior managers who are keen for more electronic material to be used. Finally there are outcomes for users, that they will be able to use the material in the future. Related to this are wider outcomes of protecting history (particularly relevant to archivists) and creating heritage by selecting what is important enough

to be kept for the future (Cameron 2008, p.177). The digital preservation literature identified a number reasons why selection is necessary (section 2.3.4) including: there are inadequate resources to preserve everything; there is an increase in volume of material so need to focus material on the most deserving of preservation; and for quality control purposes. Expected outcomes from the selection process would therefore be a focus of resources on the most valuable material and being able to manage large volumes. Based on definitions on digital preservation, the literature provides other expected outcomes: authentic, trustworthy material that is useable long term. Respondents identified preserving authenticity, enabling and maintaining access and use, and the ability to perform preservation actions as outcomes, though here they seem to be referring to preservation as much as selection.

## **5.5 Institutional factors**

This section relates findings from analysis of the practitioner interviews that centre on the wider institutional context in which they work.

### **5.5.1 Mission**

The overall mission of the parent organisation links into the policies and activities of the institution. For example, part of the mission or objectives of the parent organisation may be to preserve heritage collections and this is reflected in the preservation activities of the institution. When asked whether the mission of the institution affected their work, respondents pointed out that it outlines what is valued by the organisation or institution, in that the value of an item changes with the purpose of the institution - if it still fits the topic of interest or the story that is being told about it for example (P5) then it has value. The guidance given to selection and other activities may be indirect, depending on how the mission is written. The mission may be very high level; P9 describes the mission of the institution he works in as being written so generally as to be '*virtually meaningless*' and P16 states that the mission is so high level that it does not affect her day-to-day job directly. However it is clear through the interviews that the mission of the parent organisation or the institution acts as either a limitation by providing boundaries to selection, or as a driver to actions by the institution. The mission of the organisation or institutions may be broadly stated, for example 'curate knowledge' (P1). In this case she explained that as the mission of the parent organisation is 'format neutral' i.e. does not specify the format only the content topic, so her institution must now select digital material to fulfil its remit. A change in the mission or focus of the parent institution would drive a change in the collecting policy of the institution (P13). The collecting aims for the institution determine the content of the material to be collected; they provide input into specific collecting criteria and as the mission evolves so does the criteria. The link is particularly clear in university libraries or local institutions which have very specific collecting aims based on the specific topic areas the university teaches or researches, or the specific local area the institution serves.

### 5.5.2 Policies in practice

In addition to collecting data on policies from the literature and from interviews with experts and practitioners, digital preservation policy documents currently in use in memory institutions were examined; this section presents the findings from this. A list of institutions which are represented in this sample can be found in section 3.6.1. The policies represent the following types of institutions:

- A national museum
- Data centres and university data repositories
- Local authority archive services
- Combined national archive and library services
- National libraries
- University, state or other large library

It is unfortunate that only one museum is represented in the sample of policies as this means it is more difficult to compare between the three sectors addressed in this research. However it reflects the current lack of priority in the museums sector for digital preservation.

Analysis of the policies demonstrates that an institution may refer to its selection policy as a collection management or collection development policy, or present their preservation or selection guidelines within other wider policies or strategies. This tendency was noted in the literature review (section 2.3.6). The draft New Zealand policy (p.11) makes explicit the relationship between strategy and policy in their organisation; from the vision set out in the strategy document preservation principles are derived. These principles then lead to policies which provide a framework for implementation; the outcomes from the implementation fulfil the original vision. Strategic policies are written at a very high level in the organisation. The policy may state the purpose of the policy itself, the purpose of digital preservation in the organisation, and the preservation objectives. All policies sampled included a 'principle statement', which states that an aim of the organisation is to preserve digital material long term. Many of the high level policies are aspirational, in that they set out the aims of the organisation and give a vision of the future capabilities of the organisation. It is clear that many are acting as public statements, rather than as internal guidance, by drawing attention to intentions and assumption of responsibility for digital preservation of the organisation. It may be that the documents which describe procedures to be followed and are aimed at an internal audience are not made public, and hence were not offered to the researcher nor were available on websites. Alternatively as suggested in the practitioner interviews procedures are not made explicit in formal policy documents and so detailed procedural documents do not exist.

Table 22 in appendix 6 gives a detailed comparison of the clauses found in the policies sampled and those suggested in digital preservation policy writing guides found in the literature (see

section 2.3.6). The table spans four pages. The clauses suggested in the guides that are most common are:

- preservation objectives
- principle statement
- contextual links
- identification of content.

Clearly the final one is of most relevance to selection and nineteen out of the twenty two policies examined had a clause relating to identification of content. The two that did not are the University of Edinburgh data repository and RCAHMS; these both have specific purposes and audiences so perhaps it was not felt necessary to include a clause relating to content. The University of Edinburgh data repository's policy was created using the 'openDOAR' tool. The openDOAR service provides tools specifically designed for institutional repositories and the policy tool consists of a series of standard sentences and 'pick list' choices to produce very short policy statements for various areas, such as metadata, preservation and submission. It is very specific to the perceived needs of an institutional repository but it is unclear why it is so short.

Most of the policies sampled included a clause which identified the content for which they are responsible or to which the policy applies only in a general manner. These are general clauses with many stating that selection would be performed according to other policies and gave links to specific selection or collection policies; the Parliamentary Archives digital preservation policy states '*Appraisal and selection procedures will comply with Parliament's Records Management Policy and Collection and Acquisition Policy*' for instance. Formats are mentioned in a few policies; for example the Northumberland Collections Service (the county archives and local studies collections) states that '*Popular formats, which are supported by as wide a range of software as possible, are therefore to be preferred where possible.*' Others make general statements that they will store material in acceptable preservation formats, without explaining which. There are statements in some archive policies which make clear that appraisal of digital records will be performed using the same principles as non-digital material; for example the Greater Manchester Archivists Group policy states that '*The 'Collections Development Policies' and 'Appraisal Policies' of local authority archive services should be applied to digital born records.*' It seems that in digital preservation policies institutions give little detail about the actual material they will select and few give indications of criteria. 'Requirements for acceptance' clauses detail the requirements of the institution for the material it will accept and unsurprisingly half the data centres give this information in their policies, along with a large library that has a digital a preservation programme and a university repository. It is surprising that only three policies contain clauses relating to intellectual property when from the literature and the interviews it is clear that this is an important driver and limitation on preservation activities. Five

out of the seven policy guides also recommend that clauses relating to intellectual property are included in policies. It may be that institutions identify intellectual property with specific procedures which do not need to be detailed in a policy setting out strategy and aims. It would seem prudent however if a clause relating to intellectual property were included in order to reassure depositors and provide information for those who may wish to re-use the material.

### 5.5.3 Policy

All of the respondents were asked if they had policies relating to digital preservation or selection in their institution. A variety of policies exist according to the respondents, sometimes referred to as 'strategies', though this may also refer to documents created or applied at a higher level in the institution than a policy. Disposal or deaccessioning policies were most often linked with collection policies. Findings bear out Sinclair et al (2009, p.280) that those institutions with a policy were more likely to have plans and solutions for the long term management of digital material. Highly engaged respondents worked in institutions that were more likely to have digital preservation policies.

The majority of institutions of all types said they have a collecting or acquisitions policy, though this referred in the main to non-digital material. Respondents at twelve institutions said they had a written preservation policy for either analogue or digital, representing about half the respondents, with one claiming to have be developing a preservation policy for paper material. Seven said they had a policy specifically for digital preservation, with a further four stating that their institution was in the process of developing such a policy. Some respondents spoke about their intentions to have a preservation policy in the future, in response to the intention to collect digital material. None of the respondents reported having a policy specifically for selection of digital material. Similarly to the suggestions in the policy guides (section 2.3.6) and the findings from the analysis of extant policies in the previous section, indications of material to be selected are part of a broader collecting policy in some institutions which applies to both digital and on-digital material. P14 for example:

*'We do and we have our collections policy, our collecting policy, and of course the nature of a collecting policy is that it is medium-blind.'*

P19 put forward the argument that policies will become more important in the future as there will be more digital content; it will be more important to be clear about how this material will be managed and the resources needed. Pressure from outside organisations, such as funders, will drive the creation of preservation policies.

Policies may be created as part of a wider set of collection management policies, with more than one function or activity described in one document, such as 'acquisition and disposal' policies. P19 makes this explicit:

*'So in terms of the preservation policy how was it devised? That very much stands within the bigger picture; it wasn't a standalone preservation policy.'*

Policies for the institution may be affected by policies and aims of the parent organisation; when these change the institutional collecting policy also changes (P13).

Of particular relevance to selection is the role of the policy in providing guidance for collecting and prioritisation, with many practitioners considering it a responsibility of management to create policies for guidance purposes. One archival manager (P14) indicated that this was one of the ways he influenced how his service selected material, as he did not get to work directly as an archivist anymore. Senior managers are clearly able to steer the focus of collecting in their institution through the formulation of policy or objectives. The policy clarifies the scope of the collection and collecting activity:

*'...and then everybody who's responsible for acquiring material simply refers back to it. And kind of goes, does it fit? What does the policy say? Policy says a, b, c this doesn't fit therefore it is out.'* (P13)

Items that are considered for collection are 'tested' against the collecting policy; it becomes a measure of value to the institution -P13 stated that *'if that item whatever it is fits with the collection policy then to us it has value.'* Despite the insistence of P13 that the collecting policy is the ultimate guide in his institution, he admits that the scope of the collecting policy is very broad so it is possible for a very large range of content to be found suitable for collection within the policy. The policy is therefore *'susceptible to interpretation'* and practitioner judgement is used. This reflects the policies examined in the previous section which are high level and not very specific about what will be collected. Some respondents were keen to point out that although they did not have a formal policy they had informal ones that guided their work i.e. they had policies that are not written down. Whilst this may be useful it could also lead to unsystematic or inconsistent selection by different practitioners within the institution. Even when a policy is formally written down it may not be referred to unless necessary – it seems that procedures are often carried in respondent's minds. Policy is *'inherent in what we do'* (P7), until the practitioner wants or needs to refer to it for a particular purpose. This is a function of the professionalism of the practitioners and reflects their view of themselves as a professional. The policy allows the curator to make decisions and gives room for her to use her judgement in what to consider for acquisition (P5). This purpose reflects that suggested in TNA guidance (2011, p.5) which states that policies provide authority for those undertaking selection. This is also demonstrated by an archivist (P12) who uses the policy as a source of authority for selection decisions made at short notice:

*'... if you're in a sticky spot for whatever reason and you've got to make a decision about something or you're being challenged by a member of the public it's having a document there that you can kind of refer back to....'*

These types of policies detail the procedures to be followed rather than being aspirational or strategic and with it P12 is able to justify her decisions. P15 describes his institution as having high level stable policies, available on the website, but also practical procedural documents which, due to the rapidity of change in digital preservation, are not stable and are not necessarily made public. The practice of having both a high level document, partly for reputational purposes, and practical non-public documents for procedural guidance is a pragmatic answer to the need to address different audiences with policy.

#### **5.5.4 Legal, ownership and risk**

The respondents were asked directly which legal issues affected their work and in response they mentioned much of the legal framework that applies to their domain as discussed in the literature review. The practitioners were all aware of problems relating to copyright legislation as it affected their work, though some were more aware of detail than others, in particular the managers. If the institution had a dedicated officer who managed legal or risk issues then the practitioners had less detailed knowledge although they were aware of who to consult. The earlier interviewees, especially archive respondents, made clear that data protection and freedom of information issues in particular were important, so follow up questions relating to these two specific issues were included in later interviews. The practitioners described difficulties they have in practice in complying with legal requirements, in particular the Data Protection Act, when managing material. Respondents reported difficulties with distinguishing the content of records, similarly to issues with non-digital, such as where the subject line in an email might not be relevant to the contents so each may need to be read (P7), or that there may be more problems with distinguishing sensitive personal information from professional information where senders put both in one email. Boundaries between personal and professional may be blurred, for example in social media and email correspondence. Digital material is seen by respondents as more problematic than paper as high volumes and problems being able to access material in older formats means it is not so easy to go through every item and isolate the things that should not be there. P11 described the problem where she only has time to check a selection of the material that they have deposited and then it becomes a risk management issue, where the resources needed to check more of the material is balanced with the risk of releasing potentially sensitive material. Two archivists questioned the influence of data protection legislation on selection. They suggested that archives may become blander if the more sensitive material is not kept, either by the selector attempting to avoid problems or by the creator becoming more aware of potential for embarrassment and so being more wary of what they create and deposit. Whilst she made clear this had not happened in her institution, P20 suggested that:

*'I mean with selection with literary archives you very very quickly come up against DP issues partly because writers are so indiscreet. So there is, I'm sure there would be, a temptation of let's just lose that, you know it's going to be litigious to retain that.'*

The temptation not to select material because it may cause legal problems is an ethical issue; ethics and legal issues were often conflated by respondents, in that when asked about ethics they immediately mentioned legal issues. This could have implications - what would they do if they had to do something legal yet unethical? This was discussed by two archivist respondents (P11 and P13) in relation to a case in Australia where an archivist destroyed records at the instigation of his employer. Whilst P11 was adamant that she would adhere to professional ethical requirements, P13 questioned whether in reality this would always happen.

The most common legal issue mentioned by all respondents focused on the ownership of rights, such as copyright, and the consequential restrictions on their actions. Ideally the institution would own the material and thus be able to take whatever actions are necessary to preserve it, but it is clear from practitioners that this is not always the case. This issue was addressed by the respondents partly through discussion of deposit agreements and rights. Non-digital items may be deposited with an institution on different terms; loans of twenty five years, a permanent loan or as a gift were all mentioned. The terms of acceptance have implications in the way the institution can use the material and make it available. For example if it is gifted then the institution owns it and can do what they want with it. Problems arise for example if the institution wishes to change the material in any way, such as by copying digital material for preservation purposes or making it available on line, then permission should be gained from the legal owner. If the owner does not respond to communication then, as P12 complains, *'... so I'm stuck with this CD that doesn't work but I can't exactly throw it away because they've deposited it with me for twenty five years...'* The answer to this according to respondents is to create a deposit agreement that includes clauses to enable the institution to use the material. P20 mentions for example that her institution includes clauses in the deposit agreement that there is a presumption of permission to put limited extracts for non-commercial use on the internet.

The respondents also make it clear that there is a risk management decision to be made between the difficulties associated with not owning the item or the copyright and losing the item or access to the item. The terms of deposit may act as a limitation on the preserver's actions, potentially putting the item at risk. However some material is valuable enough to collect without ownership; P21 mentions:

*'But because the potential value of having interviews with [...] and the alternative is that they just get thrown away, that's a decision that's been made in the past that it's worth taking these things on. Even though we had no way of exploiting it without permission of the sponsor or the production company.'*

Some respondents especially from larger institutions mentioned that they have a manager responsible for assessing and managing risk, who is often the same person that manages legal compliance issues. This is unsurprising as much of the risk mentioned by respondents seems to come from complying with legal requirements. It was apparent that institutions are risk averse and this is characterised by many respondents as a risk to the reputation of the institution.

Reputational risks mentioned by the respondents included:

- Deaccessioning or selling material which they are seen as responsible for or as being valuable
- Not complying with legal requirements
- Not ensuring continued access to material such as that made available online
- Not being a 'good custodian' of material – donors will not continue to donate if they do not feel that the institution is reliable and safe
- Not providing authentic material in an archive

Risk is also addressed by many respondents in terms of the risk to material from software or hardware dependency. This causes anxiety amongst many respondents. However amongst those respondents who are highly engaged with digital material there is confidence that future change can be planned for. P13 and P15 in particular felt that format obsolescence is not the main risk, perhaps reflecting their greater familiarity with the digital preservation literature where commentators have questioned the focus on obsolescence (Rusbridge 2006; Rosenthal 2010). P15 acknowledges there is a risk from media obsolescence but in his institution they have made a decision not to accept removable media, thus negating the risk from these becoming obsolete. He goes on to say that:

*'There are an awful lot of things that are talked about in the academic literature around digital preservation that in practice are not the proximal risk that an institution is likely to suffer from. The proximate risk is actually can I get any data at all? The proximate risk is when somebody gives me a CD what do I do with it? They're not 'will Microsoft Word become on obsolete format in the next three months'? Because it ain't going to happen.'*

Similarly P13 is clear that '*obsolescence happens over time, not overnight*', so there is a '*window of opportunity*' where change can be foreseen and planned for; he feels his institution has control over the format and media on which digital material is held.

### **5.5.5 Finance and resources**

The needs of funders influence selection and prioritisation decisions in particular, whether they are internal or external funders. P7 works in an archive which serves and is funded by more than one local authority, so in order to satisfy the funders the origin of the material needs to be taken into consideration when deciding what actions to take, or which to take next. Where

institutions are funded by external partners and the institution has an explicit commitment to a project, that project may take priority over other work (P8; P19). Work that the institution is not explicitly funded for such as appraising and cataloguing large archives (P11) especially where the material is held on old or obscure media or formats may be pushed down the priority list. P19 describes it as a *'constant reshuffling of resources and priorities'*. P20 feared that the more easy to use material will take priority as funding is only short term and funders want measurable outputs in that time.

Funding for preservation is from internal as well as external sources and is often controlled at a very senior level. P19 is a middle level manager in her institution and she described balancing demands from her staff when putting together bids to senior managers for funding for projects and activities, including digitisation and preservation. They then take an overall view of organisational priorities to decide funding allocations between different areas of the organisation and between different activities, of which preservation is only one. Five practitioners, mainly from the museum domain but also from non local authority archives, mentioned that the institution has selection committees which review or approve practitioner selection decisions and consist of more senior management. This seems to be in order to control expenditure and applies most clearly to acquiring non-digital material. However P10 also stated that that the committee in his institution were interested in overall costs to acquire, which includes preservation; this may become more common. Many respondents mention their funding is being reduced now or in the near future and this puts pressure them. The greater time needed to appraise and catalogue digital material is also a concern to respondents. P19 is very clear that in order to deal with this, the workflows and resource requirements must be fully understood so that appropriate resources can be made available and the institution understands the implications of collecting material. She also cites financial considerations as a driver toward working more closely with colleagues who also have an interest in preserving digital material in order to make best use of limited resources.

Cost of an item acts as an institutional criterion. When making a decision it is not just a question of 'do we want this?' but also a question of 'what are we prepared to pay?' (P9). This reflects the comment by E1 that *'it's not just the money it's what is worth the money.'* The value of the item must be balanced against the costs of ownership, and resources to preserve material into the future are uncertain (P8). Budgets for digital and paper are not always integrated (P3) and there is no extra money for collecting and managing digital, so it has to come from existing budgets (P19). Continuing funding is very important as responsibility for preservation does not end when the funding for a project ends:

*'That's one of the problems that's happened because obviously our project stops, our funding stops, we've got to try to pull it in within everything else'* (P18)

There may be changes to collecting by practitioners due to an increased awareness of the resources needed to sustain preservation activities. It is recognised by respondents that acquiring material commits the institution to further costs in the future, whether this is for sustaining digital preservation or undertaking conservation actions on fragile non-digital material. Processing the material for preservation or taking conservation actions add to the overall cost and this must be taken into account when deciding whether to accept material. P13 describes an instance where his institution spent eighteen months working on digital material which was in an obscure format to make it readable, only to find that it was not suitable for the collection. As he puts it:

*'...the problem with digital over physical is you can't lay digital out on a table and quickly cast your eyes over it and say this is fantastic stuff or this is rubbish. If you've got obscure formats on obscure media you have to invest lots of time effort and money just to lay it on the table.'*

Some library respondents report that they can currently collect what users might become interested in but this may change as it will be expensive to preserve more digital material. In their view this may mean collecting becomes more focussed on 'now', immediate use and demand, rather than collecting 'just in case', which will happen less. They also identify that decisions need to be made to decide between acquiring access to different resources, asking questions such as 'which is the best value for money?' (P16) and that expenditure needs to be justified according to use statistics (P17). The cost of processing material to make it accessible must also be taken into account when selecting material, similarly to the process for non-digital material. P19:

*'So when we take in donations, physical donations, we're now sort of looking at them ... How much would it cost to make them accessible through cataloguing, collection care etc etc? So we're looking at that sort of front loading end which is quite key and I think you know those work flows will follow through into how we deal with digital content.'*

With non-digital archival material there is a cost associated with providing appropriate ongoing storage with the right environment ('passive' preservation). Similarly digital is a 'treadmill' that will require constant management and expenditure; P22 hopes that it will get cheaper and easier to store and maintain digital material, or if not cheaper then perhaps more cost effective. Time is also a resource identified by respondents that influences what digital material is kept, as it takes longer to appraise each piece of data, so it is easier to keep more material rather than taking the time to appraise each individual file. P7 explained that in her experience this means more 'rubbish' is kept. Only one respondent (P10) specifically mentioned the time pressure created through the need to take action sooner for digital than traditional material due to the fragility of the medium.

The influence of the market on preservation was discussed quite emphatically by one archive manager who insisted that if the material is valuable enough then there will be an economic driver for preserving it. He argued that material that is in a more pervasive format has a greater chance of being preserved due to the influence of economic scale – if enough people use a format it will continue to be supported and tools will be developed to read it; this reflects the position of Rosenthal (2010). Therefore a single institution probably does not need to worry about developing such a tool itself, even if it could afford to do so (P15). This contrasts with other respondents, who perhaps do not have the strategic overview that P15 has, and so worry more about commercial formats becoming obsolete. Some respondents criticise commercial companies when describing their dealings with them, which include a commercial record storage company, publishers who pay for digitisation projects in order to exploit the images and major electronics companies. P5 describes how one of the latter in particular refuse to work with her museum, as the company views museums as being concerned with ‘old’ heritage, whereas they want to be associated with the ‘new’. Many respondents were very clear that their institution is not a commercial enterprise even though it may make money from exploiting assets, such as through marketing activities. P9 describes items in his institution as having archival value but also an ‘image’ value, where image value is ‘...for reproduction, for sales, for commercial, for wow factor’. Despite the overall respondent eagerness to distance themselves from commercial activities and describe the more difficult aspects of working with companies, one institution has developed a digital preservation service where they commercially exploit their expertise. In contrast to other respondents, P19, who is a part of this institution, questions whether institutions should change their ‘*psyche of collecting*’ to include more commercial activities.

## 5.6 Conclusions

This chapter has presented findings from analysing the interviews with practitioners working in a variety of libraries, archives and museums. Personal factors which influence selection by practitioners were identified. Their insistence that selection is a professional activity, which requires special training and specific skills, was very noticeable, and they view their responsibility for preservation to continue with digital material. Some respondents seem to view this responsibility as particular to them, and the concept of sharing responsibility between individuals seems to be a source of some anxiety. Sharing responsibility seems more acceptable in institutional collaborations, although here there is also some anxiety. Sharing responsibility seems to be seen as appropriate for infrastructure and resources but not selection. This implies that practitioners wish to retain a professional identity which includes selection, and to keep control.

Stakeholders with an influence on selection have been identified, including managers, senior managers, users, creators and donors, funders, other organisations and IT staff. These are similar to stakeholders in selection for non-digital material, with the exception of IT staff.

Practitioners were particularly concerned with the relationship with IT as this is new and there is the perception that IT and curatorial staff do not understand each other's perspective. This seems to be an area which warrants both further research (Oliver, Chawner and Lui (2011, p.314) note that little has been done) and attention from managers to negotiate such relationships so that selection is performed effectively for the institution. Mandates from senior managers are clearly necessary to enable selection and preservation to take place; difficult relationships with other stakeholders can be alleviated or made worse by senior management. Senior managers and subordinate managers also act as limiters on selection through the setting of policy and finance. External funders may also provide limitations through the conditions under which they are funding selection and preservation, and also indirectly through pressure to reprioritise preservation activity.

Other colleagues include those working more closely with the practitioner as well as those in the wider professional community. Colleagues are used as 'guides' to value and help in selection decision making and both formal and informal networks of colleagues as guides are important. This suggests that building relationships within the institution and with external organisations that have greater expertise is a more necessary than for non-digital material. Practitioners also identified many sources of material; these can affect selection through formal agreements and by providing information on material that may need urgent action taking to save it. As a lack of resources implies that collecting 'just in case' users may want digital material is less likely, although practitioners clearly still do this as they do not know what future users will want, respondents were clear that their decisions are user led. However it is not clear how this happens more than for non-digital material considering their insistence that selection is a professional responsibility.

It is clear that practitioners do not see a direct role for non-professionals in selection as they are seen as too biased or subjective, and lacking the necessary 'historical sense' skill in particular. According to practitioners this skill is developed through training, and it allows them to take a wide view of what should be preserved for the future, differentiating them from non-professionals. Although professional judgement is seen as inherently subjective, it can be used and trained for by professionals. A range of skills, knowledge and particularly personal attributes were suggested by respondents as being important in those who carry out selection activities. This indicates that although practitioners view themselves as being distinct because of their skills and training, it is also a distinct set of attitudes which allow them to perform selection subjectively, yet on behalf of others. The necessary knowledge, skills and attributes suggested by practitioners are mainly generic; they are the same as for non-digital material, with the exception of a greater emphasis on 'keeping up to date'. This suggests that practitioners already have many of the required skills, but that extra knowledge of developments specifically in digital preservation is necessary. Three types of skill sets were identified by respondents: only having curatorial skills but working with technical experts; developing a digital preservation

or curation specialism, similar to those specialists who currently work with manuscripts; and becoming a generalist where the practitioner also has technical skills. This indicates that the role of 'digital curator' is developing, though this may depend on the type of institution; a lone worker is more likely to need a generalist set of skills whereas someone working in a large organisation either works with IT staff or has the opportunity to develop specialist skills.

Engagement with digital material clearly affects how confident they are about selecting digital material. There is a perception that digital material and preservation is changing at a rapid rate. Engagement was demonstrated as a range with low, medium and highly engaged respondents. Their reactions to digital material range from disinterest for the lowest engaged respondents, through anxiety, to complete confidence for those that work daily with digital material. Factors that influence the level of anxiety shown by respondents include: having clear lines of responsibility between stakeholders and in managers; having clear workflows, in particular for how to select or appraise digital material when it arrives in the institution, for controlling sensitive information, for managing rights issues and how to select or appraise large volumes of material. The latter suggests that practitioners are seeking guidance on how to select specifically digital material and as yet that have not found it, neither in their institutional policies and procedures nor from the wider digital preservation community. This is apparent in the methods suggested by practitioners to improve selection guidance, including using stricter criteria, having clearer policies to guide selection and developing strategies for sampling material. It was noticeable that the concept of significant properties, important in the literature and to the experts, was shown to be of much less relevance to practitioners, even to those who understood it and were highly engaged. This implies that there is a disconnection between digital preservation theory and the practical needs of practitioners.

The concept of value was deliberately explored with respondents, following the review of the literature and the expert orientation interviews. Financial value was clearly of secondary interest to most of the respondents, although they did discuss value in terms of costs of preservation or potential for commercial exploitations, similarly to the experts. The concept of 'history' was raised by respondents in terms of historical value, both to the present and the future. Respondents focused on 'use' value to the institution and in the case of archivists on archival value as a record. That value changes and can be influenced was acknowledged by the respondents; however they were unclear as to how it could be measured. Methods of determining value suggested were comparative, including comparing it to the collecting policy, the current collection and to collections held elsewhere.

The selection process as described by practitioners is similar to that described in the literature, with three stages: becoming aware, deciding value and action. Selection happens not only at initial collection or acquisition but also later in the management of the material. Strategies for selection are easily identifiable for web based resources, including bulk or domain harvesting

and an event based selective approach. The criteria-based approach also seen in selecting web resources is assumed by practitioners to be the strategy that will be used to select other digital material, with little change from traditional selection. Criteria suggested were similar to those used for non-digital material although it was clear that practical criteria relating to the ability of the institution to preserve material are highly influential.

Institutional factors that influence selection were explored with the practitioners. The legal environment is an important factor influencing selection; in particular the management of rights which either enable or limit the activities practitioners are able to do. Legal and ethical issues are clearly closely linked - when asked about ethical issues respondents often replied by discussing adherence with legal requirements, although some also mentioned adhering to professional codes and concepts such as impartiality. This conflation of legal and ethical issues could have implications for what is selected; for example what are ethical actions, such as avoiding censorship, may not be legal in some circumstances.

Finance and costs along with commercial interests of both external funders and the institution lead to a reshuffling of priorities for selection and preservation. The resources of the institution, including staff, space, storage and time are strongly influential as limitations to selection. Finally the mission of the institution provides the direction and framework for broad collecting priorities and acts as a driver to actions. These factors provide some of the boundaries to selection.

Policy has also been a focus of this chapter. The practitioners described the types of policies that they have in their institutions, revealing a variety of policy provision. Those institutions with a range of policies, including for preservation, seemed more likely to be engaged with preserving digital material, suggesting that a policy can act as a driver to further engagement with digital material. Policies were often created by finding other policies and using them as a template or by consulting with colleagues. It is clear that collection policies are useful tools for practitioners as they control collecting activities, although some respondents suggested that not all policies are formal. Informal policies may mean that the reason for selection decisions are unclear, but allows the practitioners to adapt their practices more flexibly. This then allows them to use their professional expertise and have some control over their decisions. This is similar to the use of criteria, which may also be unstated. Findings from an examination of current digital preservation policies have also been presented in this chapter. Many of the policies obtained were high level strategy documents, especially those from national institutions. It became clear that the terminology used in labelling documents as policies or strategies can be confusing; this is unhelpful to those using the policies. These documents seemed often to act as a public declaration of intent or as a vision document from which further more detailed policy could be created. Detailed policies documenting practical or procedural clauses seem more likely to not be made public; this allows the institution more flexibility in their response to changes in best practice and removes the need for policies to be rewritten each time there is a change. In

comparing the policies with the policy guides examined in the literature review, the most often recommended clauses were most often found in the policy documents, implying that many institutions are following best practice guidelines. Most of the policies had a clause identifying the type of content for which the institution is responsible, even if this was only described in broad terms with little detail and few criteria.

The next chapter brings together the findings from this chapter and chapter four, discussing them in relation to the findings from the review of the literature and the research objectives.

## CHAPTER 6 DISCUSSION

---

This chapter synthesises the findings from the analysis of the expert interviews presented in chapter four and those from the practitioner interviews in chapter five, along with the review of the literature presented in chapter two. This chapter will highlight key issues and concepts derived from the empirical data, relating to the original research objectives. Those identified throughout provide a deeper understanding of selection for digital preservation in different contexts. A conceptual model of key underpinning factors influencing selection has been devised based on key factors and concepts, in order to fulfil objective 5: *To construct a framework of factors in the practice of selection for digital preservation and make recommendations for future selection activity and research*. This will be presented and discussed in this chapter; recommendations will follow in chapter 7.

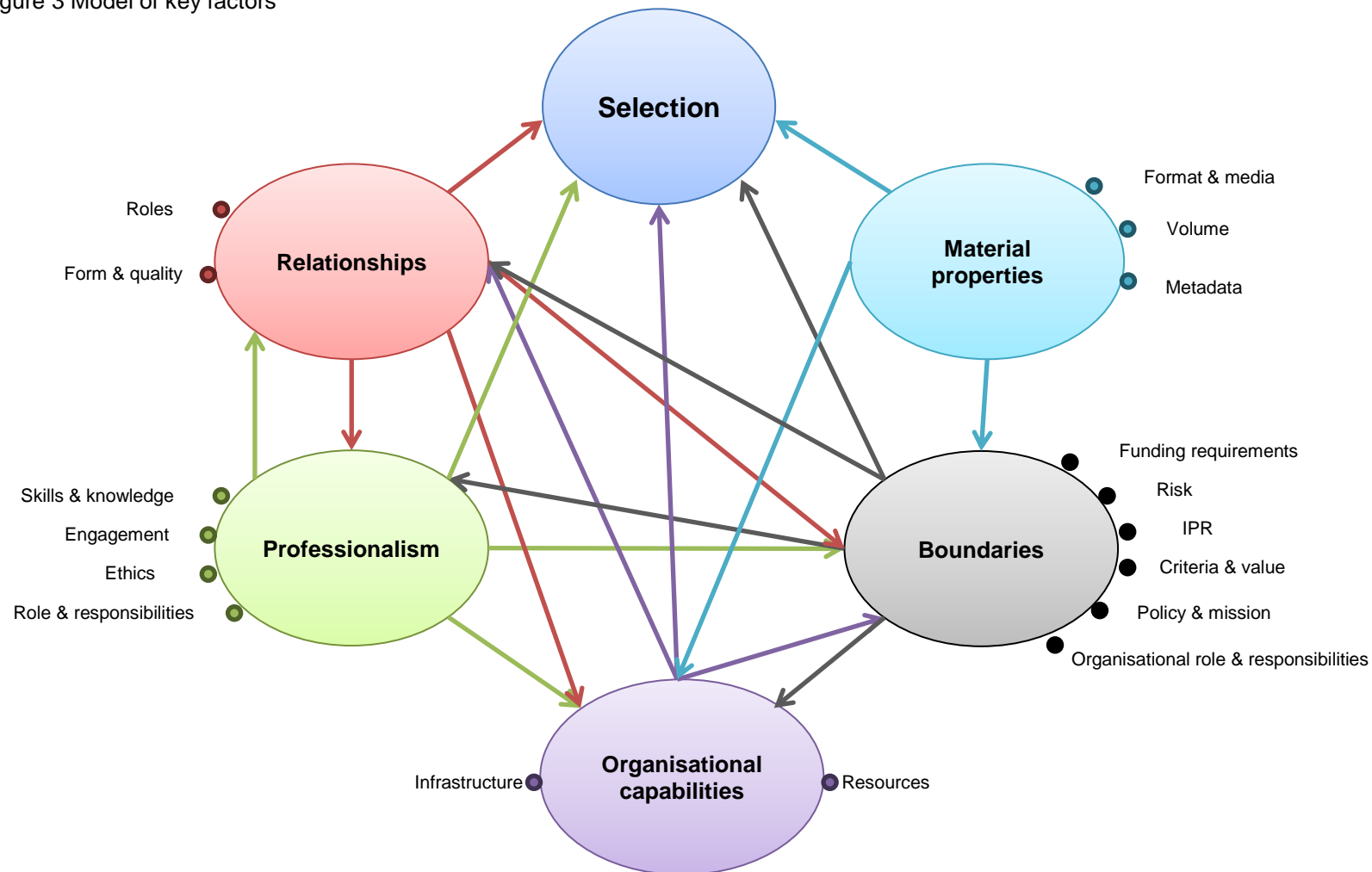
## 6.1 Conceptual model of key factors

Key factors in selection are the focus of objective 4: *Identify and describe the key factors in selection for digital preservation in memory institutions* and objective 5: *Construct a conceptual model of key factors and their relationships which influence selection*. Through analysis of the data gathered in this research, and distillation of the key factors identified through objective 4, a model of key underpinning concepts in selection was developed and is shown in Figure 3. The model brings together concepts of relationships and professionalism; the concept of boundaries comprising of factors which inhibit or drive selection; and concepts relating to technical issues, of material properties and organisational capabilities. It illustrates the inter-relationships between key concepts, with arrows representing influence, and the key factors within these concepts are also shown on the figure as small circles. These factors are summarised separately for extra clarity in table 17, and will be discussed in the context of the wider conceptual model in the rest of this chapter.

Table 17 Key factors

Concept	Factor
Relationships	Role
	Form and quality
Professionalism	Skills and knowledge
	Engagement
	Ethics
Boundaries	Role and responsibilities
	Criteria and value
	Legal and funding requirements
	Rights ownership
	Policy and mission
Capabilities	Risk
	Resources
	Infrastructure
Material	Format and media
	Volume
	Metadata

Figure 3 Model of key factors



### 6.1.1 Boundaries

The concept of boundaries used in this research refers to those influential factors that provide structure and boundaries to selection as well as those that allow or drive selection, defining the scope of selection. These are often contextual factors, such as policy or legal requirements, that are outside the direct control of the practitioner. Selection for non-digital material also has limitations but for digital material these seem to be particularly important. Having clear boundaries helps both practitioners and other stakeholders to control selection. The findings show that limitations are experienced particularly as rights issues. The ownership of rights was cited as an important influence on selection throughout this research, reflecting findings in the literature review (see section 2.3.5). Evidence of an inability to select or deselect material because rights were not owned or because the rights owners were untraceable were found in the practitioner interviews. There was also some evidence of copyright ownership being an example of an unstated criterion, where there is a disinclination to select something that the institution did not own or could not acquire the rights to. It is clearly important that rights are clarified and negotiated, and are influencing and being influenced by the form and quality of relationships with creators and donors, similarly to non-digital. The findings suggest that formal deposit agreements or access arrangements assigning responsibilities for preservation should be agreed when something is acquired as this means later preservation actions can be taken without having to seek the owner's permission. Surprisingly little was mentioned regarding the preservation of e-journals however; it was expected from the literature review that this might be an issue for library respondents but concerns regarding e-journals focused more on access. There was an assumption that issues of preservation responsibility were being dealt with at a higher level, reflecting the findings from the PARSE.Insight project (2009), Durrant (2008), and Meddings (2011) who reported a desire amongst their respondents that national institutions have a role in preserving e-journals. This emphasises the leadership and advocacy role held by stakeholders such as (other) national institutions and by organisations such as JISC.

The concept of risk has been highlighted in this research, where the focus of risk to the material and risk to the institution follows that found in the literature review (section 2.3.5). The risks identified in the interviews are similar to those for non-digital material. There is risk to the material of loss at each selection point and risks to material from obsolescence were recognised. However the evidence suggests that with appropriate engagement and planning the latter is now not as acute; supporting the arguments from Rosenthal (2010) that format obsolescence is not the most pressing risk. Other risks relating to selection were identified in this research, including a risk from not collecting contextual information which could affect the ability to find or use the material and determine authenticity; risk to the institution from not preserving material either to fulfil legal requirements and provide evidence of activities or in not choosing to preserve material it is responsible for which could damage its reputation. These are similar to risks found for non-digital material and reflect those found in the literature, such as from Vermaaten, Lavoie and Caplan (2012) and Clifton (2005).

The concept of value clearly underpins selection and different types of value were identified in this research. Financial value is of little interest in the cultural heritage context, although the experts identified 'opportunity costs', or what Whyte and Wilson (2010) referred to as 'non-replicability', where the cost to reproduce it is too great. Findings suggest that material that has a current or potential use is more likely to be selected, or if it is linked to the function and aims of the institution or the professional. The concept of place is clearly still relevant to digital material as although access to it does not depend on geography, the aims of institutions are still linked to places and communities that they serve. Questions might therefore include:

- Does the practitioner or colleagues want to use it, for an exhibition for example?
- Does it fulfil a user need?
- What can it do that will be of benefit? For example does it add to the current collection, tell a story, or record an event?
- Is there deferred value i.e. will it be useful in the future to increase understanding or represent a recent event, function, person or something else now? Will keeping it reduce future risks?
- Will it bring financial or other benefits to the institution or to users in the future?

The findings imply that whilst understanding different types of value is useful, identifying a justifiable use for material is particularly important. This will vary according to context; for example the first question in the list above may be more relevant in a museum. There seems little point in preserving material if it is not or will not be useful in some way; digital material does not have value as an artefact in the same way as non-digital material, although the hardware needed to run it might. Value lies in what it can do or be used for and here this research confirms the view of Beagrie et al (2011, p.56) where it is the outcome of preservation that is important, not the preserved information. This is linked to concepts of history and the future, highlighted in this research, where outcomes of selection include 'keeping history safe' and creating heritage i.e. choosing material to be passed to the future.

Measuring value in this context is clearly an inexact science due to its subjective nature, and few formal methods of determining value were used by practitioners. Assessing value is subjective and criteria for non-digital material varies between institutions, however core criteria for digital preservation were identified in this research as: technical ability and infrastructure, costs, the issue of intellectual property rights, collecting policy or mission, the existence of documentation, with the evidence indicating that the first three of these are the most influential. Criteria clearly limit selection and are considered further in section 6.3.1. Similarly policies also act as limitations on selection and these are discussed in section 6.3.2.

### 6.1.2 Capability and properties

The concept of capabilities refers to those factors which affect the ability of the institution to select and preserve material; this is clearly linked to the properties of the material. This includes factors relating to not only the availability of appropriate infrastructure but also the provision of appropriate resources to enable selection and preservation. Unless the institution has the requisite capabilities or envisions developing them in the near future there seems little point in selecting digital material. The issue of resources has been widely examined in the literature (section 2.3.4) and the importance of adequate financial provision for sustainable preservation was similarly highlighted by both experts and practitioners. Those resources identified in this research are similar to those needed for non-digital material - finance, time and space for example. It is clear that selection requires expertise and time; there is a danger that material will be selected because it is easier to select at a larger scale than taking the time to look in more detail. There is a link in the model from capabilities to relationships, indicating where a lack of capability may prompt greater collaboration with others or a realisation that guidance is needed from other institutions to improve capabilities. Relationships with funders are of importance here also as they can provide the means by which institutional capability is improved.

Properties and boundaries are linked through a consideration of criteria used to select digital material. Format affects even highly engaged practitioners as they use it as a criterion for selection. This confirms the suggestion in the DPC Decision Tree (2006) that a question to be asked should be '*is it in a format you can manage now or in the future?*' Whilst most of the practitioners were keen to state they did not select on the basis of format they clearly did; a common criterion for non-acceptance was that the information was stored in unsuitable formats which their institution is unable to process or preserve.

This research has highlighted a difference between digital preservation theory and practice. The review of the literature (section 2.2.2) and the experts identified significant properties as an important issue. The importance assigned to it was not reflected in the responses from most practitioners; as noted in section 5.3.2, many practitioners rejected or lacked familiarity with the concept. Those practitioners who were highly engaged with digital preservation were more likely to be familiar with this concept. The digital preservation literature is written by experts and academics who may be removed from the day to day concerns of practitioners, or the term used may not reflect the language used by practitioners, so there may have been some misunderstanding. However this implies that there is a disparity between theory and practice.

### 6.1.3 Professionalism

The concept of professionalism refers to the role and responsibilities of practitioners, their skills and their ethical behaviour as a professional. It was noticeable that ethical behaviour was often equated by practitioners with adherence to legal requirements; however it is clear that these are

not always the same. What is legal may not be ethical, such as choosing not to preserve records embarrassing someone important or not choosing to preserve material that you disagree with politically. Findings show that this is an issue in archives as appraisal procedures are based on theoretical principles and is not related to the context in which they work; P11 for example was clear that her loyalty is not to her employer but to her profession. It is clear from the literature review that librarians also have a professional ethic to select or appraise without bias (IFLA 2012; CILIP 2012b) although there is no legal requirement for them to do so. By considering ethics mainly in terms of legal requirements it is clear that this could have implications for selection practice and professionalism; it is also clear throughout this research that this does not change with a change in media.

Respondents clearly gain confidence from their professional training and it seems to be part of their professional identity that selection is their responsibility. It is clear is that the majority of skills and qualities needed by practitioners will be very similar to those for non-digital material and those found in this research to be necessary are largely generic. Apart from technical knowledge it is only the ability to keep up to date that was regarded as different for digital, presumably relating to the perception that digital formats and media change rapidly. Three skill sets were evident in this research: individuals with separate curatorial and technical skills, as seen in institutions where practitioners and IT staff are working together; that practitioners will have to develop technical skills and become generalists; or that a digital curator specialism or profession will develop. This research has revealed that this role is developing; in some highly engaged institutions there are individuals with titles similar to this and responsibilities which include both selection and technical aspects. Unsurprisingly, this was suggested by the digital preservation experts as a desirable development. It is not clear however how technical this role will become. The question arises of the feasibility and desirability of the structure of professions being changed so that 'digital curator' is either a separate role or specialism, such as conservators or those who can read medieval Latin are now, or that librarians, archivists and curators should be expected to multitask and deal with digital in a generalist role. As more material is being created digitally, practitioners will generally need to update their knowledge and skills to encompass 'digital preservation' in order to be able to understand the implications of choices that they make. Those who do not may be professionally disadvantaged; this may increase over time as more material is in a digital format. Practitioners may need to consider appropriate levels of expertise, as with traditional material; a high level of expertise is not necessarily needed in order to accept and use parchment but you do need a great deal of expertise in order to undertake conservation work on it. The traditional model of a conservator being consulted when necessary could be applied to digital material; expertise can always be bought in from elsewhere or be found elsewhere in the organisation. This is where collaborations and personal relationships are particularly valuable in order to access the required expertise; practitioners should create and develop interpersonal contacts themselves. This would enable many practitioners to continue working as generalists with the support of

digital preservation experts when necessary. The digital curator role may become more prevalent and although some evidence of this role was found in this research, it is also clear that co-operative working to share skills is more common. This is consistent with findings from Verheul (2006, p. 29) who found that none of the libraries sampled had one department responsible for digital preservation but instead they worked together. The evidence from this research indicates that a combination of skill sets and roles will be appropriate which depend on the context the practitioner works in. The range and depth of skills required to work alone is different to that needed for working in a large organisation, where it is more likely that specialisms can be developed. The danger is that a digital curator role may focus on technical skills and the curatorial skills will become less important; traditional skills, such as reading Latin, may be confined to even fewer specialists, though these skills will continue to be needed.

This research has revealed that practitioners conceptualise digital material in different ways, as 'just stuff' or as something 'different'. Those that are highly engaged with digital material are more likely to view digital as a continuation of non-digital; they do not regard it as very different to their existing responsibilities and are positive about preserving it. They acknowledge that there are technical differences, but were comfortable and knowledgeable about it; they see the commonalities with non-digital material. The majority of practitioners however are aware of the problems but as yet do not have the skills, knowledge, policies, or procedures, to integrate digital material into their normal work, hence digital is seen as something different and this creates anxiety. The findings revealed that digital material worries practitioners because they do not know what to do with it; there are not the appropriate workflows in place to care for it appropriately. This is both in the short term before it is considered for permanent acceptance and in the long term, especially when it is an unexpected acquisition or in an unfamiliar format; and they lack of confidence in their understanding of technology. The answer suggested to this from the findings of this research is to increase the engagement of staff through training and exposure to networks of expertise and by developing guidelines and workflows that integrate digital with non-digital, providing reassurance to practitioners. In this way it becomes less 'different' and more 'just stuff'. By doing this it is more likely that digital material of value will be selected as practitioners will be more confident in making decisions.

Professional boundaries between practitioners and other stakeholders are undermined with digital as it requires technical knowledge that may not be held by practitioners, and this clearly causes anxiety. It is therefore in the interests of the practitioner to reinforce professional boundaries; practitioners are clear that responsibility for selection is theirs and that through their training they have a particular ability to select. The practitioners were very sure that valuing material is a professional task, dependent on judgement developed through training and experience. They claimed in particular a 'historical sense' skill. This allows them to judge more accurately what contemporary material will be useful in the future, as well as judging what historical material is important now. This is similar to the observation by Shen et al (2011, p.16)

that users search for titles to satisfy current needs whereas librarians select with a view of future needs.

#### **6.1.4 Relationships**

The concept of relationships describes the range, quality and form of relationships that influence selection, from informal personal relationships to formal collaborations between institutions. It underpins factors relating to the role of other stakeholders, which is discussed in section 6.2, the quality of those relationships (whether they are useful or hinder selection) and the forms interaction may take between stakeholders. This concept has been emphasised throughout the interviews with experts and practitioners and this research clearly supports the argument from Lavoie and Dempsey (2004) that selection for digital preservation is a social and cultural process. A key point from the findings is that negative relationships engenders a feeling of anxiety in practitioners from a loss of control over selection, so it is important to create harmonious relationships to reduce anxiety and encourage effective co-operation.

Whilst different types of interactions may be seen with non-digital material, the nature of some relationships is different for digital. Key forms that interactions take which have been identified in this research include 'being told what to do', where the practitioner has little choice about what to select, such as where an archive has to preserve the records of its parent organisation. In this case material is not initially selected by the archivist and it was clear from the practitioner interviews that there can be difficulties collecting such material in digital form. This form of relationship could also be seen where the practitioner does not have sole control of selection and preservation. IT or systems staff have a role which is largely unacknowledged in the literature, as noted by Oliver, Chawner and Lui (2011, p.314). However this is clearly an important relationship specifically for digital material and the quality of the relationship can influence selection. For example P4 described an instance where the IT department in her organisation was a distinct hindrance to selection. Relationships can have an effect on the boundaries that surround selection; unclear roles and poor quality of relationships with IT staff for example mean that the practitioner may have less control over the material and workflows to process the material may not be adequate for the needs of the practitioner. In contrast, positive relationships with sources can increase selection through increasing the availability of material.

A second form apparent in this research is 'collaborating', where selection decisions are made jointly with another stakeholder. This is seen formally between institutions, though in this research formal collaborations were uncommon; many institutions seemed to be developing their own systems and workflows for digital material and in one instance, despite a formal collaboration to share resources, an institution was making sure that they kept their own. Findings from this research seem to indicate that despite the evidence of institutions collaborating on digitisation projects from Portico (2011, p.15) for example, there is much less, especially on a non-national level, for digital preservation. This is explained as a matter of trust,

by both experts and practitioners, which does not yet seem to be in place despite recognition of the benefits of collaboration. Organisational capability and the limitations to selection are clearly affected by this factor; relationships with external stakeholders can influence the resources and expertise available to the institution for selection and preservation activities. Building useful relationships through collaborations with other institutions or third party service providers can improve the technical infrastructure of the institution, improving capabilities, and this is a key criterion for selection.

Another form of relationship apparent in this research is where the practitioner is able to influence others to help with selection through making appropriate choices. There is a greater imperative with selecting digital material to control format and gather contextual data and this research has provided evidence that the practitioner is able to influence the creator to choose particular formats. It is clear that relationships with creators and donors also allow practitioners to gather contextual data and that this should be part of the acceptance procedure; if the relationship with the creator or donor is poor or non-existent then this will influence the amount and value of the contextual material gathered and so the value. Whilst this is not new for digital material it seems more likely to happen, especially in archives, where contemporary material from accessible creators or donors may be collected or deposited. This seems to cause an ethical dilemma for archive practitioners which has not yet been resolved as it means that it is possible to influence the creation of the archive, undermining the traditional view of appraisal and the archivist's role that positions them at the end of the life of the records. Evidence from this research suggests that the way forward is to provide clear selection criteria and procedures, identified as boundaries in the model. This will make sure both creators and practitioners understand the boundaries of selection, and help to apply these consistently, improving both relationships and the exercise of professionalism of the practitioner.

A further form of relationship apparent in this research is 'seeking guidance', where the practitioner looks for guidance from others to help decision making. This is a clear example of relationships improving professionalism as skills and knowledge are gained by the practitioner through using others as guides. Practitioners were very clear how important and valuable formal and informal networks of expertise are to their selection practice. Immediate colleagues are particularly important, and their influence on the practitioner performing selection is high. Immediate colleagues have been shown to provide extra guidance and reassurance about difficult selection decisions. There is clearly a greater use of support networks with digital material and these should be further developed and made more accessible.

## **6.2 Stakeholders**

Objective 2 in this research is: *Identify internal and external stakeholders in selection for digital preservation in memory institutions*. Questions within this objective ask: who are the

stakeholders? What are their roles and responsibilities? This section synthesises and discusses findings presented in previous chapters relating to these questions.

Through this research a range of stakeholders with an influence on selection has been mapped and their roles and responsibilities in selection traced. These include:

- Practitioners
- Colleagues
- Managers
- Senior managers
- Institutions
- Users
- Creators and donors
- IT staff
- External funders
- Other institutions

As noted in the digital preservation literature (section 2.3.3), digital preservation is a shared responsibility between different stakeholders. Evidence from this research suggests that although responsibility for selection sits firmly with the practitioners, other stakeholders have an influence depending on their role. Findings regarding the responsibilities of others that impact on selection are summarised in figure 4 on p.174. This represents all domains as issues of stakeholder responsibility were often found to be similar. Stakeholders at progressively further distance from the job of selecting are shown in a series of concentric circles, with other stakeholders represented in boxes with dotted arrows, representing potential or indirect responsibility.

The issue of responsibilities is an important one; the BRTF report (2010, p.43) suggests taking responsibility for preserving material implies an institutional commitment to providing a large number of resources over time. There are stakeholder responsibilities within the institution that can affect selection. Within the context of selection the institution has broad responsibilities, including caring for collections (regardless of media), working with stakeholders, having clear policies and procedures, providing support for staff and working within legal requirements. These provide the overall structure for selection and as figure 4 illustrates, the institutional responsibilities are reflected in the responsibilities of internal stakeholders. The next layer toward the centre of figure 4 represents senior managers. This research has highlighted the importance of senior management 'buy in' for selection of digital material and the need for them to provide a mandate for collecting digital material, along with resources, which allows selection to happen. Examples were found in the practitioner interviews of respondents who wanted to engage further with digital material being limited in their activities due to a lack of senior

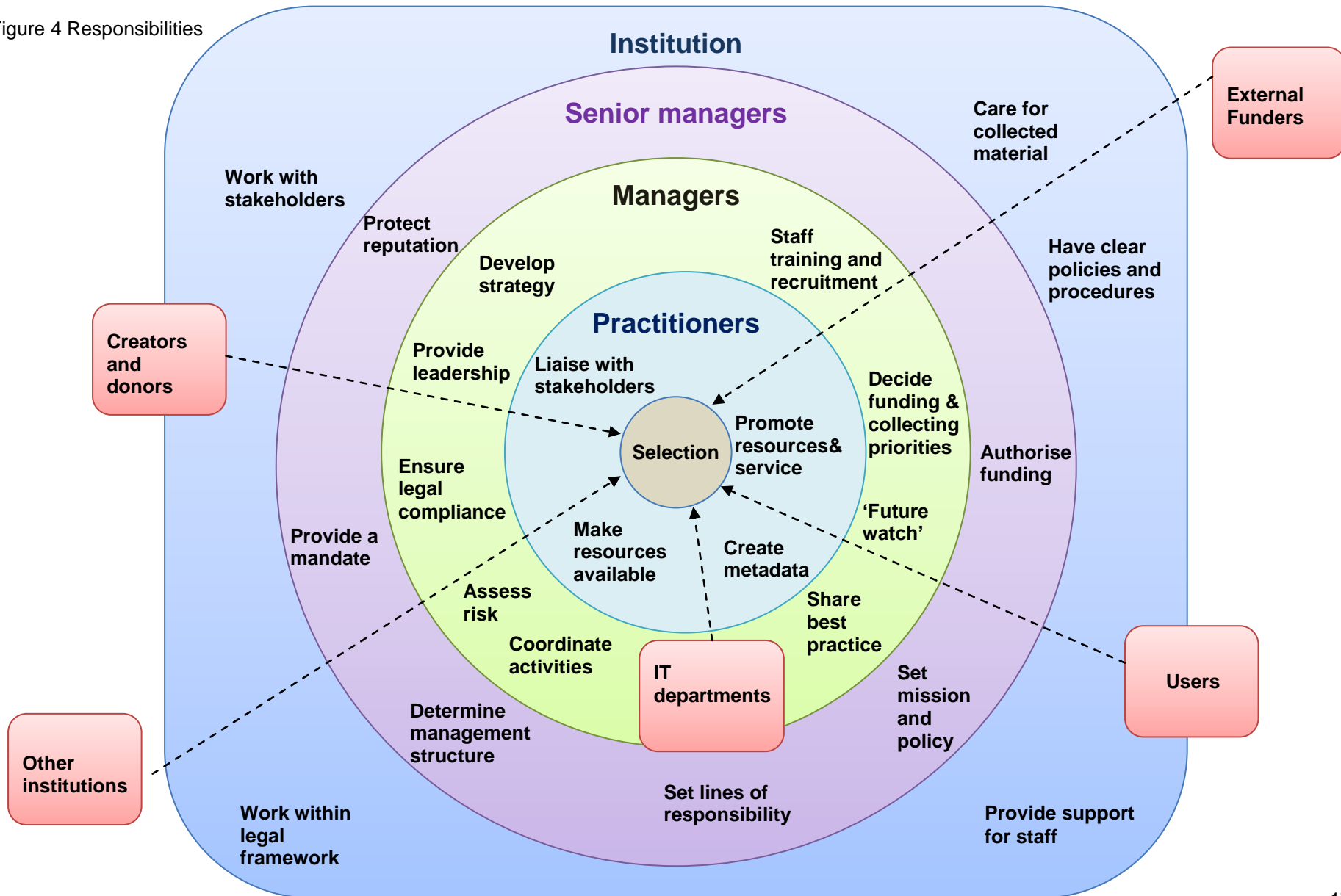
management support ( P4, P12, P14 for example). These findings are consistent with conclusions from Hudson (2012, p.46) and the DPC Handbook (DPC 2008, p.20). It is clear that without institutional leadership and 'championing' for digital, the efforts of individuals will have little impact.

The next circle towards the middle of figure 4 represents lower levels of management, some of whom took part in the practitioner interviews. Findings again indicate a generic set of responsibilities; these are not related to media, and reflect good management practice within an institution. The one responsibility which is highlighted for digital material is 'future watch', referring to keeping up to date with developments in technology and digital preservation practice. This is included as one of the stages in digital curation in the 'Lifecycle model' (Higgins 2008), though there it is termed as 'community watch'. There is clearly a need to keep up to date with digital preservation issues by managers, which they would then be able to pass on to practitioners through their responsibility to 'share best practice', or by arranging for them to attend appropriate training courses or events. Making sure that staff have access to the appropriate training and there are adequately trained staff within the institution is also a management responsibility, which is generic but perhaps more pressing with digital material as the evidence indicates that there is a link between engagement and knowledge of digital preservation and the ability to confidently select for it. The inner circle nearest selection represents practitioners; selection is clearly part of their view of themselves as a professional and this has not changed from non-digital. The responsibilities found in this research for practitioners are again not related to media or domain but are generic. In comparing between domains it is at a more detailed level that differences begin to appear; 'make resources available' for example, may be for display in a museum, online for library users or searchable in an archive, and selection itself is clearly performed in different ways for different purposes. However the clear message from examination of the evidence relating to responsibilities from the expert and practitioner interviews is how little these will change specifically for digital material. The authority of the memory institutions and the practitioner in selecting material continues to be assumed throughout the research. The evidence clearly suggests that the question of 'who has responsibility for selection of digital material?' has the answer 'the same people who do it for traditional material'.

Other stakeholders with an influence on selection are included in the figure 4; users are shown as being both internal to the institution and external and these were both shown to be important, as are creators and donors who can also be internal and external to the institution. The research indicates that a responsibility of creators and donors seems to be to make selection easier for practitioners by bringing material to their attention and using appropriate formats; this is discussed further in section 6.3.3 in consideration of relationships. Other institutions and external funders have also been shown as having influential responsibilities, through controlling resources to select or through taking responsibility for preserving material. IT staff were clearly

shown in this research to have an influence on selection through their responsibilities to manage the technical infrastructure that selected material is stored on, and in some institutions to manage the material itself. In figure 4 IT staff are within the institution, and as yet none of the practitioners had experience of outsourcing IT provision, indicating that this is still rare. This may be due to the lack of trust of third parties, as suggested by E2.

Figure 4 Responsibilities



Findings from this research indicate the various roles in selection that other stakeholders have: users, guides, limiters, collaborators and sources. These roles are not exclusive; users can be sources, for example, and they can have a direct role in selection, such as through self-selecting what to deposit in an archive or repository. Although these roles are similar to those with traditional selection, the evidence suggests that the social aspect of stakeholders as guides is more important as practitioners need more support in selecting digital material. Guides are those colleagues that the practitioner turns to for formal or informal advice and so they are influential when making selection decisions. Whilst this is similar to advice given by colleagues for non-digital material, it seems that practitioners may be more reliant on guides through not having the requisite technical knowledge to make fully informed decisions. Although networks of digital preservation expertise exist, through the efforts of the DPC and DCC for example, this research demonstrates that the reach of such networks is not far enough. There is a lack of awareness of sources of guidance, a lack of opportunity to consult colleagues and a lack of management support to attend training or awareness raising sessions. Practitioners strongly valued being able to get advice from others who are engaged with selecting digital material and preservation, but it was clear that not everyone could or had accessed such support.

The evidence suggests that users are either current customers of the institution, or conceptualisations by practitioners of potential customers in the future. By conceptualising a future user, the practitioner is able to select for the future using current guidelines. Selection in libraries in particular focuses on knowledge of the user as central to selection decision making (Clayton and Gorman 2001, p.4; Harvey 2005, p.58; Johnson 2009, p.108) and practitioners claim this is only possible through skills developed in their training; they imply that they are in a position of the 'expert' in relation to others who have not had their training, lending support to the assertion by Cameron (2008, p.180) and Smith (2006, p.12) that heritage is circumscribed by 'experts'. Though the experts, in particular E8, are clear about problems of not representing or including different views, this role is seen as largely unproblematic by the practitioners. The evidence suggests that the scope for users to be collaborators in selection is circumscribed, and that the traditional role of a practitioner in acting as a 'proxy', or stand-in, for the users when selecting material does not change with digital.

Evidence found through this research suggests that the role of collaborators may become more important with a greater likelihood of more collaborative working being necessary for digital material. This research has underlined expected benefits to institutions of greater collaboration through sharing resources, such as described by Angevaere et al (2012, p.95); Lunghi et al (2012, p.214), but it also highlights the lack of attention paid to informal collaboration between individual stakeholders. Despite the acknowledgement of the potential benefits of formal collaboration, there is a reluctance to engage in this with other stakeholders for digital material, especially IT staff, without clear boundaries in responsibility due to a perceived loss of control. The benefits of formal collaboration between individuals are not always obvious to them. These

relationships need careful and deliberate management, in order to build trust and a shared understanding, such as through arranged meetings where all parties are able to contribute. Here the potential influence of senior managers is apparent in providing a mandate and a requirement for different areas of the institution to work together. Findings suggest that most training is done 'on the job', underlining the importance of informal methods. Informal collaborations through networks seem to be an effective method for spreading awareness and knowledge, and staff should be encouraged to participate.

By making clear the aim or mission of the institution, devising policy and controlling budgets, senior managers limit selection for both digital and non-digital material. Throughout this research it has become clear that other stakeholders limit selection, including creators and donors, IT staff, external funders and other institutions. Practitioners are limited in their choices of what to preserve not only through formal boundaries set by others, such as policies, but also in more indirect ways such as through the ability of others to affect the resources available to select material. External funders are shown to be particularly influential as they affect the priority given to certain materials for preservation as well as the resources available. Findings indicate that sources of material have an influence on selection as firstly they determine in part the 'anti-collection' that a selection is made from, identifying material to be preserved, which echoes a suggested role of private individuals from the BRTF report (2010, p.39). Secondly sources have been shown to affect the value of material, for example creators and donors may limit selection through their choice of format, where use of an obscure format means it is less likely that the material will be selected as the institution may not be able to manage them. The quality or quantity of contextual information sources supply also clearly affects whether the material is of value (stated by E4 and P10 for example), supporting the suggested criterion of 'documentation' from the DPC Decision Tree (2006). Also value can be affected through the ownership and management of rights by sources; value has been shown to be closely linked to the use of material, consistent with the argument from Ross (2007, p.3) that high value may be equated with high use, so it is of less value if the institution is unable to use it now or to preserve it for later use.

### **6.3 Practice**

This section relates in particular to objective 3: *Investigate the practice of selection for digital preservation in different UK memory institutions through the examination of practitioner views.*

Within this objective were a number of research questions:

- How is digital material selected for preservation in those institutions which have a digital preservation function?
- What drivers and barriers are there to selection in different memory institutions?
- How do selection policies relate to selection practice?

- What assumptions have been made by memory institutions in their current thinking about selection?
- What changes to practice do practitioners perceive are necessary to select digital material for preservation?

This section will discuss the findings relating to these questions.

As preservation is an on-going management process (DPC 2008 p.24; ALA 2009), selection for it can happen at different stages, not only when selecting for particular preservation actions, and these selection points have been clarified. The evidence clearly suggests that selection of digital material in institutions which are managing digital material already is very similar to that for non-digital material, in particular when the selector actively identifies material to acquire or to preserve. It is clearly desirable to have a positive relationship in order to exercise greater control over what is donated or deposited, which is possible for those working in institutions which serve a limited community that is easy to contact, such as an institutional repository. Those practitioners who are able to have a positive relationship with creators and donors are able to influence them in their choices, of format in particular, which makes selection easier.

There seems to be more difficulty and much greater anxiety with selection of material that is donated or deposited especially if it is unexpected; there is less control over the format it is in and the amount of contextual information supplied. The findings suggest that not only do many practitioners not know how to store or process digital material when they receive it, the volume of files received is greater than that through non-digital donations or deposits. P22 mentioned having hard drives with terabytes of photographs in his institution for example. Though this problem may be more acute in archives, the evidence suggests that it is applicable in museums and libraries also as they accept donation or deposits. The effect of volume on selection was clearly demonstrated; without looking at each digital file on a CD or memory stick for example it is not always obvious what it contains, especially if the donor has not supplied adequate metadata. The lists of criteria suggested in the literature, such as Whyte and Wilson (2010) assume that this is already clear, but this may not be so. This suggests a preliminary question should be 'do we know what this is?' The next question is then whether it is worth preserving and the answer may be 'no', even though there may be valuable files contained, as there is a need to balance the potential value of the content with the cost in time and resources to determine the content and to create adequate metadata.

### **6.3.1 Criteria**

This research has identified a wide range of criteria used in practice in different contexts. Although many criteria identified in this research as being used in practice are clearly based on those used for non-digital material, differences were found. One difference between criteria for

traditional material and for digital was in the priority given to different criteria, with a higher priority given to:

- The ability of the institution to continue to provide access
- The ability of the institution to read, store, and preserve the material
- Format or medium
- Costs to preserve
- Rights ownership
- Duplicate material

There was agreement amongst respondents that adhering to collecting policies is important but the evidence also suggests that unless the institution has the technical ability to use and manage the material, there seems little point in comparing it to the collecting policy to determine the value to the institution. Similarly value is of secondary importance to the provision of adequate resources as even if the material is valuable there may not be the funds to preserve it or make it available. Practitioners were definite that format was not a criterion; especially in archives where practitioners were clear that they should collect material regardless of media or format. But through analysis of the interviews it became apparent that format is used as an unacknowledged criterion for both highly engaged institutions already preserving material and those which are not yet at that point, as suggested in the expert interviews. It seems clear from the evidence that the criteria found in the literature and listed above are rightly of higher priority, once the preliminary question of identifying content has been answered.

In chapter 2 a table of core criteria was formulated and this is similar to the list of high priority criteria above. In combining the findings from all the stages of this research the core set of primary questions are:

- Technical ability and format - is the institution able to read and manage the format that the material is currently in? Does the institution have the ability to store and take appropriate actions to preserve the material? If not, can the institution access these services from elsewhere?
- Costs - can the institution afford to process, preserve and continue to provide access to the material for as long as necessary?

This contrasts with the priority suggested in the DPC Decision Tree (2006) which begins with an assessment of value and comparison with the collecting policy. The second section then turns to acquiring or negotiating intellectual property rights and findings from this research support this. Suggested questions that need to be answered here include: can intellectual property rights be acquired or agreed in order for the institution to preserve and use the material? If not, can these be acquired at a later date? Does the risk of losing the material outweigh the risk of

not currently owning the rights? The Decision Tree then asks question relating to technical ability and finally documentation and metadata, a factor which is supported by this research as being of high importance.

Further informal or unstated criteria were also identified in this research. These include having enough time to appraise and comparing potential items to those collected by other organisations. Lists of criteria, such as those developed by the Significance 2.0 framework or Whyte and Wilson (2010), do not take account of not only the differences in institutions and domains but also in the knowledge and skills of individuals, and so are only be useful as general guides and the evidence from this research suggests that they are used rarely. The practitioners were comfortable with using both formal and informal criteria but this could plainly cause difficulties, as there would be no clear justification for collecting something and if the practitioner left then they might take that knowledge with them. The 'ad hoc' nature of criteria may lead to problems of consistency in the application of criteria when selecting material. This means that whilst practitioners use their judgement it is still clearly important for institutions to clarify their selection criteria in order to provide guidance, as argued by Seadle(2004), supporting the recommendation from McInnes and Phillips (2009) that policies are updated to contain selection criteria for digital material.

### **6.3.2 Drivers and barriers**

The research has demonstrated that drivers to selection have an influence on selection itself. Selection is driven by a number of factors which were noted in the literature, such as value (Tanner 2012) and risk (section 2.3.5). Economic drivers were emphasised by the respondents in this research supporting the suggestion, for example from Harvey (2007, p.8), that costs of preservation and availability of finance drives the need to select material that is most worthwhile. This research also provides evidence that economic drivers affect both the outcome of selection, in terms of what can be afforded, and also when the selection is done. Examples of these include the need for easy 'wins' following demands from some funders and their influence on the prioritisation of selection activities; and an increased awareness of costs to preserve digital resources, so practitioners feel less driven to select 'just in case' their users may want material (although it was clear that this practice continued, especially in archives).

The evidence also suggests that preserving institutional reputations is a driver to selection, with E3 focusing on the reputation of an institution as a 'good custodian'. These findings support Hughes (2012, p.6), who suggested that value in material may come from adding to the reputation of the institution. What also became apparent in this research is 'opportunity' as a driver; selection is not planned for but performed when there is an opportunity. This is a concern for digital material as preservation should be a managed activity (DPC 2008) or it may become unusable or of little value. The expected 'pressure of time' driver expressed in the literature (Rothenberg 1999, p.2) for example) in which digital material needs to be selected and

preserved quickly is not yet apparent. This could be due to the recognition by highly engaged respondents that preservation and material change can be planned for and also from acceptance, demonstrated in particular by museum respondents, that not everything of value can be saved. Whilst 'imminent risk to material' is mentioned as a driver in all domains, this was in the context of material being destroyed rather than material becoming unusable. Risk to the material seems greater from organisational factors such as not having a planned workflow or sustainable funding, than from material degradation. This reflects and supports the condition for sustainable preservation suggested in the BRTF report (2010, p.78) of 'appropriate organization and governance of preservation activities'.

Barriers to selection were identified through this research and they focus on institutional and individual factors. It is clear that the responsibility for preservation needs to be recognised at a high level within the institution; if this is not recognised then, as this research demonstrated, any efforts by practitioners to select digital material are in vain as there is a lack of funds and a lack of infrastructure to manage the material. This reflects the reluctance noted in the literature of institutions to take responsibility when incentives are unclear (Currall and McKinney 2006). In addition factors relating to individuals are important barriers, such as a lack of knowledge of sources, of appropriate appraisal and selection strategies when confronted with digital material and of what to do with digital media when it is acquired, as they prevent the practitioner from making decisions. There is a lack of confidence in practitioners relating to digital which affects their decision making which is not recognised in the literature. This research suggests a number of solutions; this could be ameliorated through the institution having a set of policies and workflows detailing how selection or appraisal should be performed and how the material should be managed. Clear criteria need to be created for the practitioners and procedures for storing and accessing digital material whilst selection is performed; this would increase the confidence of practitioners. Also institutional agreement of who is responsible for on-going management of material would help to clarify workflows. There needs to be leadership and management support for preservation within the organisation in order to emphasise the need for selection; clear support structures and a management 'champion' to take forward issues such as not receiving deposits or a lack of co-operation from other stakeholders. The evidence suggests attention to these factors would do much to reduce internal institutional and individual barriers to selection.

### **6.3.3 Policies**

The review of the literature revealed a number of guides to digital preservation policies, such as Beagrie et al (2008), though much fewer specifically for selection such as the DPC Decision Tree (2006). From analysis of the guides key digital preservation policy clauses were identified. These are:

- A principle statement
- A statement of preservation objectives
- Indication of preservation strategies
- Contextual links with other policies or documents
- Identification of content
- Standards compliance

The majority of policies which were then analysed were found to include these clauses, along with general expressions of what is collected and preserved. It became apparent through analysis of extant preservation policies that the most common strategy for policy content is to not include detailed criteria in the publically available policy but to keep these as internal guidelines. There is a clear benefit gained by the institution from having both a high level policy, which gives a general outline of what the institution collects and preserves, and a more specific policy or procedural guide with detailed criteria. Whilst this is practical, as the public document does not need to be revised often, there is also evidence from this research that it is of benefit to make public more detailed criteria regarding what the institution will accept responsibility for. Practitioners clearly find these particularly useful to guide donors, enabling them to control the expectations of other stakeholders about selection more easily. The findings show that they also value policies very highly to guide their selection, where the value of material and likelihood of selection is increased if it meets the collecting policy. Having a preservation policy is also clearly of benefit to the reputation to an institution with preservation responsibility, supporting the suggestion from TNA (2011b, p.6) that it is worthwhile for an institution to create an aspirational policy, or, as suggested by the experts, a policy which refers to an external preservation service. It seems there needs to be a balance between the clear benefits of having more detailed policies publically available with the amount of resources it takes to keep them up to date.

#### **6.3.4 Assumptions and change**

What is surprising in this research is the lack of change envisioned by both experts and practitioners; archivists especially insist that underlying archival principles do not need to change to take account of different media. Evidence shows that many skills are generic so will not need to change, although specific technical skills will. A number of other aspects of selection seem to remain the same, such as being able to make a judgement about the value of material, and high level policies which guide and limit selection based on the mission or focus of an institution. There is an on-going assumption that practitioners are the best people to select on behalf of users which continues for digital, in part due to their 'historical sense' skill, and findings suggest that the underlying responsibility of collecting institutions to care for their collections does not change; figure 4 in section 6.1 demonstrates that responsibilities are largely generic.

The analysis reveals a clear view of digital as being the same as non-digital material, it is 'just stuff', once the practitioner has been able to extend their knowledge and become engaged with digital material. The anxiety and negative feelings engendered by digital are ameliorated by acceptance of preserving for digital material as a continuation of traditional practices and responsibilities. There was also a message from experts and some practitioners that change can be planned for; highly engaged practitioners were especially positive. Whilst this is not new, practitioners that are not in touch with digital preservation literature or training seem to be less aware of this. A further assumption is that criteria used in selecting non-digital material is appropriate for digital and the evidence from this research supports the view that many criteria are used for both digital and non-digital, with the exception of criteria based on the institutional ability to process and preserve, as discussed in section 6.2.1.

Despite these similarities, some changes to practice were identified by the respondents. An important change is that of relationships with other stakeholders, as more stakeholders become involved in the preservation of digital material (see section 6.1). The role of the practitioner will need to adapt to include working with groups, such as IT, that they may have had little experience of. There was resistance among some practitioners to changes brought by the need to include other stakeholders and also the possibilities brought by digital material to include others, such as users, in activities that they felt were part of their professional role. There may be a danger in viewing digital as the same as non-digital that opportunities for including other stakeholder groups and using the material in new and innovative ways may be missed. The development of a 'digital curator' role has been suggested in this research and there is an assumption by the experts that a digital curator role is inevitable. Evidence from the practitioners suggests that whilst this role is developing within some institutions there are other options and it seems more likely that the range of skills and skill sets will become evident.

Practitioners will need at least some awareness of digital preservation to manage donations for instance, which are increasingly in a digital format. The medium engagement group of practitioners especially were aware of this need for change in their knowledge and skills to take account of digital material. The findings suggest that there is a need to increase engagement with digital material though greater access to training and events. This will raise awareness of not only the problems of digital material, which many respondents were aware of, but also the potential solutions, so that practical guidance and workflows can be implemented in an institution. Again, senior management support is clearly essential to support staff development and the evidence suggests that in some institutions priorities will need to change to take account of requirements for further training. Larger institutions clearly have an advantage as they have the capacity to allow staff the time to train and the resources to send them on courses or to provide in-house training. Here the ability to access networks of expertise is important; those who work in smaller institutions or are lone workers will need extra support. Findings also suggest that changes to practice could be driven however from the influence of external factors,

such as a restriction of resources or from developments in professional guidance. Whilst this is not new for digital, it seems to be more acute, with some practitioners expressing an awareness of the greater impact of resources for example, of costs to process material prior to selection, and the need to justify expenditure.

The interviews with respondents have also demonstrated areas which are as yet unknown or neglected. There is an assumption in the literature that selection or appraisal processes will need to change for digital material (Russell, 1999; Eastwood 2004, p.202; Bailey 2008, p.72 for example). The literature suggested that selection may need to change due to increased volumes, a greater number of duplicates or versions, and a greater need to process material prior to appraisal (Paradigm 2008). The influence of volume and the need to process material prior to selection have been demonstrated in this research, and 'duplicated material' is clearly an important criterion for practitioners. But there have been very few suggestions as to how selection and appraisal might be usefully changed to take account of these and the evidence suggests that up to now there have been few practical changes to selection or appraisal procedures from those outlined for non-digital in chapter 2. Strategies used to select from large volumes of web resources for example, are clearly based on existing strategies in archives; evidence from archival practitioners suggest sampling and collecting information about particular events are already used for non-digital material. Only one respondent suggested that methods of demonstrating authenticity may need to evolve to include procedures and documenting changes and provenance; this is similar to views found in the digital preservation literature where authenticity is demonstrated through using metadata to trace changes (DPC 2008, p.24; Salza et al 2012, p.26). There was also a suggestion that greater contact with a creator would enable the practitioner to have more control over the material. A further suggestion was that stricter criteria and policies should be devised and implemented. It was not made clear by respondents that these suggestions have been applied or who should develop them further. The respondents also did not seem to consider changes to practice in relation to value, where value may be added by the practitioner. Whilst it was recognised that value can be influenced, such as by highlighting material to users so they may use it more, the increased potential for 'adding value' to digital material through curation activities was not discussed. This seems short-sighted on the part of practitioners and is an example of how the view of digital material as being the same as non-digital might be limiting as potential uses of digital material are not always recognised. There seems to be a reluctance to acknowledge their own role of practitioners in 'creating' heritage through selection, which is important for digital as serendipitous preservation outside of institutions may be less likely than for non-digital.

#### **6.4 Summary of key findings**

This research had the overall aim of investigating the theory and practice of selection for digital preservation in UK memory institutions. The findings relating to this aim are as follows:

- The concept of professionalism is important to practitioners and underpins many of their activities. It includes the exercise of subjective judgements, developing a 'historical sense' skill, having appropriate domain-specific knowledge and skills, and undertaking specific training.
- The roles and responsibilities of the practitioner as the selector and their personal requirement to be a good professional, including exhibiting legal and ethical behaviour, is an influential factor on selection.
- Practitioners conceptualise digital in different ways dependent on their level of engagement. Those who are highly engaged with digital material and preservation view digital as 'just stuff', similar to non-digital material, so it does not cause anxiety and is preserved as a normal part of business. However many practitioners view digital as something different and this provokes anxiety.
- There are three different skill set types for digital material: that similar to non-digital i.e. the expertise is bought in when necessary, such as with a conservator; that practitioners would have to develop technical skills themselves; and the 'digital curator' profession will develop alongside traditional domain specific roles. This research has found evidence that this is developing.
- Research findings indicate that a review of existing availability and access to training for digital preservation would be beneficial for those involved in selection. It would enable practitioners to select digital material with more confidence and increase engagement. This has a number of aspects; knowing about appropriate training opportunities, being able to access training through being allowed to attend courses or events, gaining access to the networks of expertise, and gaining awareness of positive role models and exemplars. In particular encouraging staff performing selection to access networks of expertise seems beneficial.
- Important stakeholders that were identified include: managers, senior managers, users, creators and donors, funders, other organisations and IT staff. This differs from non-digital material in a number of ways: the influence of IT staff with whom there is a lack of understanding and common ground for practitioners; the strong need for senior managers to act as 'champions' of digital preservation and to recognise the responsibility the institution has for digital material; and the ability and desirability of having a relationship with creators which affects the contextual information that can be obtained. Without this the material is of less value and less likely to be selected.
- Roles and responsibilities of stakeholders are key factors; five different roles were identified for other stakeholders: guides, limiters, collaborators, users and sources; each role has different influences on selection. Relationships with other stakeholders are of great importance because:
  - Selection may need to be shared with other stakeholders and this causes anxiety in practitioners due to a perceived loss of control, although the

development of shared services does not necessarily mean that selection also has to be shared.

- Relationships are critical to gather contextual information, similarly to non-digital selection, but the quality of those relationships is important, including with sources not just creators.
- Practitioners require more support with selecting digital material, especially until clear workflows and institutional policy has been devised therefore formal and informal networks are more important to provide that support.
- It is not only preservation costs, the resources required and how much funding there is but also who is providing the funding for the preservation activity and their expectations which are influential factors.
- Selection is influenced by the perceived value now or in the future and the criteria used to articulate these values. These values vary by domain and are inherently subjective. Key criteria which apply to selection of digital material have been identified in this research; the primary criteria are costs, technical ability and intellectual property rights. Unstated and informal criteria are important in selection and this continues for digital material. Types of value which apply to digital material include historical value of the information it contains and archival value i.e. its value as a record. Financial value has been shown to be less important for digital material than for non-digital, although commercial considerations in terms of ways to exploit digital material are apparent. Value linked to place is still important to those institutions that are linked to a particular area; there is a desire to keep control of material, though this seems to be linked to a control of rights. Value is relative and subjective decision making by practitioners is shown to be a common approach; the use of formal frameworks is low.
- The format the material is in and the media on which it is stored can act as criteria, though these are often unstated. These factors also affect the costs to manage and the ability to use and preserve the material; this then affects its value.
- The presence of a digital preservation policy influences selection where it provides guidance to practitioners and in the limits it places on selection through articulation of collecting priorities. The high level mission of the institution is also of influence as it can drive the need to engage with digital material. There is a clear need for institutions to develop policies to guide their staff in decision making, and to reassure donors which may encourage donations through an increase in trust. This research underlines the importance of policies which has been widely discussed in the literature.
- This research has identified drivers to selection. Drivers are similar to those for non-digital material, and across domains, but of particular importance for digital are drivers from external funders who in providing funding for specific projects drive selection activity. Risk to material from imminent loss seems less of a driver than for digitisation of non-digital material.

- Barriers to selection have been identified, including: a lack of confidence in practitioners in how to select, appraise or manage digital material and a lack of guidance on this; a lack of training in selection specifically for digital material and more generally in understanding its characteristics and requirements; and a lack of senior management buy in and support.
- The issues relating to digital material are similar across domains and there are many aspects to factors which are similar to those with non-digital material, such as personal qualities needed to select, the requirement for management support, the roles of stakeholders and the limitations set by the policy and mission of the institution.
- The assumption that practitioners are responsible for selection does not change with digital. The preservation role for digital material may be taken by a new role of 'digital curators' within institutions.
- Key skills for selection do not change for digital, except for technical knowledge and the ability to keep up to date.
- The theoretical basis for selection for digital preservation as found in the literature is based on archival principles. The differences in terminology between domains and in the literature is problematic; 'significant properties' as a concept is not useful in day to day work. Whilst the research on digital preservation focuses on practical issues, theoretical concepts related to selection derived from research are not useful in selection practice.
- Key principles underpinning selection in each domain do not need to change as they are linked to the purpose of that domain, which continues regardless of media.
- It is assumed by those working in institutions that commitments made through selecting material are very long term but this is not necessarily so. The assumption of permanent responsibility in selection should be questioned.

## CHAPTER 7 CONCLUSIONS

---

This final chapter will review the original aims and objectives in relation to the findings from the research. It will contain a discussion of the main outcomes of the research, along with recommendations arising from the findings and a consideration of the contribution of this research. Suggestions for further work are provided, and then there is a reflection on the research process and considerations of the limitations of the research.

## **7.1 Research objectives and main findings**

The overall aim of this research was to investigate the theory and practice of selection for digital preservation in UK memory institutions. The research was initially based on the observation that selection for digital preservation is performed, not conceptualised, and has not been explored in the literature except in practical ways, such as by suggesting criteria. The objectives which contribute to the fulfilment of this aim have been achieved firstly through a review of the literature, which provided an overview of the theory of selection for traditional non-digital material in each domain, and an understanding of selection for digital material. An intensive, qualitative approach based on critical realism has been successfully used to undertake a set of exploratory interviews with digital preservation experts. These provided a broad overview of selection for digital preservation, building on the issues found in the extensive review of the literature and identifying influential factors in selection. Further interviews with practitioners from libraries, archives and museums built on findings from the previous two stages by examining the views of those performing selection. This was complemented by an analysis of a sample of digital preservation policies. The philosophy of critical realism that underpins this research focuses the researcher on uncovering unseen structures and mechanisms that drive observable behaviour. This focus has been usefully applied in this research, resulting in an analysis of the data gathered which enabled key factors underpinning selection for digital preservation to be identified and a conceptual model of these factors to be developed.

### **7.1.1 Existing theory**

Objective one was to 'achieve an overview and understanding of the theory of selection for preservation within UK memory institutions'. A number of key issues were identified in this stage of the research. Selection, when considered in the digital preservation literature, is seen as mainly a practical issue. Concepts used in the digital preservation literature such as significant properties and designated community were shown as potentially problematic but are widely used; these terms have specific meanings but their common use in other ways in selection, such as to signify value, may lead to confusion. Findings from this stage indicated that the views of practitioners are underreported, as is their experience of selecting digital material. Responsibility for selection is not always clear both internally and between different external stakeholders such as libraries, users and publishers, although it has traditionally been the responsibility of practitioners as 'experts' on behalf of users. The continuation of this responsibility has been questioned but not examined.

This chapter found that context was an underlying theme; it was possible to discern the differences between selection theory and practice in each domain for non-digital material. The literature demonstrated that the museum and library domains do not have the same theoretical basis for selection that underpins archival appraisal, instead basing practice on professional

principles and knowledge of contextual factors. Context as a theme was also apparent where the aims of different institutions within domains determine different selection activities; there may be differences due to scope and purpose for example an academic library has different collecting priorities than a public library. It is also clear from the review of the literature that the roles of practitioners in selection are different between each domain for non-digital material. Traditional museum selection has been ad hoc and based on the connoisseurship of the curator who has been able to use their own judgement. Appraisal in archives in contrast is based on theory which traditionally states that the archivist should not be involved in appraisal but that this is the responsibility of the creator, although more up to date archival theory recognises the role of the archivists in creating the archive. The traditional records lifecycle model of archives at the end of a record keeping process seems inadequate to describe the processes by which archivists become involved with digital material. Greater attention should be paid to the records continuum model (Upward 1996) in which the boundaries between records and archives are blurred and archivists are able to be involved earlier. In contrast, librarians have selected based on professional principles, in which they are tasked to select without bias for their users, whilst having knowledge of their collection and the institution's aims. However whilst differences in approaches were clear for non-digital material, it became apparent that there are a great many similar issues and factors between domains for digital material which transcend domains, such as copyright and the need to manage high volumes.

Selection for digital preservation is often addressed in the literature through discussion of criteria and a range of suggested criteria for selection were found. Whilst these differ by context, potential core criteria relating to digital preservation were determined which are similar to those used for traditional selection. Practical questions of technical infrastructure become more prominent with digital however. Core criteria included not only infrastructure and ability to process and preserve the material but also costs, intellectual property rights and format. Core digital preservation policy clauses were also found through the literature review which included 'identification of content' to be selected, though at this stage it was unclear whether recommended clauses were found in practice.

Terminology between different domains and between the literature and in practice seems at some points confused and this was evident in the literature review. Significant properties is a clear example of a concept which, although focused on by the digital preservation literature and the experts who are very familiar with digital preservation theory, is not understood by the majority of practitioners and amongst those who did understand it, it was felt to be largely irrelevant.

The importance of selection of digital material for preservation was recognised in the literature. Selection is an important process which has implications for the management of digital resources. However as yet there has been little tracing and conceptualising underlying

influences. It became clear that further research was required to investigate and conceptualise the influences on the selection of digital material, especially from the perspective of managers and practitioners who are involved in this activity. The selection of digital cultural heritage material for preservation in memory institutions is a cultural practice, performed within a framework of contemporary social, ethical and professional conventions. The digital preservation literature did not demonstrate a broad consideration of the social framework of selection but this research has uncovered the relationships and professional context in which selection takes place. This links to the next objective which focuses on tracing those with an interest in selection.

### **7.1.2 Stakeholders**

Objective two was to 'identify internal and external stakeholders in selection for digital preservation in memory institutions'. This objective was necessary to understand selection in a wider context than just the individual who performs it or the institutional context in which it is performed. Whilst the review of the literature demonstrated that some stakeholders are recognised and that selection is a social activity, their influence on selection was overlooked and this objective was formulated to address this.

A range of stakeholders involved in selection has been revealed, including senior managers, colleagues and IT staff, who are more numerous and wide ranging than for selection of non-digital material. It became clear that the quality and form of relationships practitioners had with other stakeholders and the roles that they played in selection were important factors. This implies that selection is a social activity and it should be viewed within a broader context than the individual. Roles of stakeholders have become apparent with sources of material, guides and limitations on selection, collaborators and users being identified. Other stakeholders have an influential role as sources of material and the relationship that practitioners have with sources can affect the material from which a selection will be made and the contextual material that can be gathered. Drivers and barriers to selection include personal, social and conceptual factors as well as institutional and material ones. Barriers to selection clearly have a social component; a lack of knowledge, confidence and engagement with digital on the part of practitioners, along with a lack of management support, were found to be important. The social nature of selection means that by encouraging relationships and networks the ability and engagement of practitioners to select digital material effectively could be increased.

The role and influence of the individual practitioner had been overlooked in the literature. They clearly see selection as their responsibility and as part of their role as a professional. However there is a disparity between their desire to retain selecting as a professional role and the ability or requirement for other stakeholders to have an input. Although practitioners are anxious when confronted with a potential loss of control over selection, it seems that by clearly defining roles and responsibilities in collaborative relationships, such as shared services, practitioners would

be able to share responsibility, though this should be carefully managed. Practitioners in memory institutions clearly need to work more closely with other stakeholders and differences in approach need to be managed, especially between technical and curatorial staff. With greater technical skills and engagement working together with technical and non-curatorial staff might be easier due to a shared understanding.

The role of 'digital curator' is developing and this implies that traditional roles may be less focused on the preservation of digital material and more on collection, use and access. There is an implication that this role could supersede those of librarians, archivists and curators for preserving digital material although this research has shown that the digital curator role is developing as one of range of skill sets. Applying existing models of the role of the conservator for example would enable practitioners to understand their relationship the 'digital curator' and integrate it with other roles. It may also allow practitioners to retain control of selection as a professional activity, even if preservation is controlled by a digital curator. The skills and personal qualities identified by practitioners are in the main the same as for non-digital material and the evidence suggests it is technical training or awareness that is required in order to be able to select digital material. The evidence suggests that professional training enables practitioners to develop an awareness of history and the future, which they termed 'historical sense', derived from their training and from experience. This broader view of the purpose of the institution and the needs of the user is clearly useful to select for both current and future users. The ability to look forward seems to be particularly useful in selecting digital material and provides a reason for selection to be performed by professionals as 'proxies' for others who do not yet exist.

The concept of engagement is important to understand the experience of practitioners with digital and their selection and preservation activities; the concept describes their familiarity and comfort with selecting and preserving digital material. Digital preservation theory is archival in nature and it is in archives that much of the engagement can be seen, whether these are archives that stand alone or are part of a library or museum. Digital material causes anxiety as it is at first seen as 'other' and as practitioners become more engaged with material this view shifts towards 'it's just stuff'. This later conceptualisation leads to an integration of digital and non-digital from the perspective of the practitioner enabling them to evaluate and select material more effectively, viewing the institutional collection as a whole. The research has demonstrated that practitioners need to become further engaged with digital material to build confidence and skills for selecting it; they require more support in order to select digital material. Networks of expertise, whether formal or informal, are key ways for practitioners to expand their knowledge about digital preservation. Therefore these should be expanded and practitioners should be encouraged to participate.

### 7.1.3 Practice

The research then focused on how selection is performed, with objective 3 being 'investigate the practice of selection for digital preservation in different UK memory institutions through the examination of practitioner views'. This objective is concerned with the practice of selection, exploring how selection is done in different types of institution through examining the views of practitioners. These views have been overlooked in the literature and the practice of selection needed to be clarified. This section also included examining preservation policies to determine if the recommendations in the literature were found in practice. This not only adds to the understanding of practice in each domain and identifies key factors, but also facilitates comparisons between selection for digital and non-digital and between domains.

The processes described though the literature review are reflected in the findings from the practitioner interviews. Selection is performed at different stages – before or after acquisition, for prioritisation for preservation actions, for specific preservation actions and for deaccessioning or disposal. The assumption in the literature that selection processes will have to change is clearly a concern for practitioners but few suggestions for how this may happen were made. This research suggests there will be little change; criteria-based selection will continue to be used but the criteria and policies controlling this will need to become more strictly defined and applied. The creation of preservation policies, including their purpose and suggested clauses, was examined in the literature review, through an examination of the policy guides and policies in practice. Common clauses and differences between policies and guides were identified, including a clause regarding identification of content although in practice this seems to be vague, limiting the practical application. Institutions need to be clear what their policy documents are for; one policy may not satisfy the needs of different audiences. If more than one document is developed then clear linkages between documents should be made. By making available their policies on preservation, institutions are able to clearly and easily enhance their reputation and practitioners use them to not only guide their own selection but also to control deposition. This research has found that the guidance does seem to be followed closely for policies, yet guidance on what goes into policies is more often gained by copying a policy from elsewhere. Increased awareness and availability of exemplar policies and good practice is required.

Selection is a subjective non-linear process in which the individual practitioner uses both formal and informal, overt and covert, criteria and value measures. Practitioners carry around in their heads criteria and knowledge that they use for selection that is informal and often based on experience. This means that their selection decisions may not always be clear or justifiable. Policy and professional ethics indicate that format should not be used a criterion for selection but the use of this is demonstrated by practitioners for pragmatic reasons. Selection is seen by practitioners as a professional activity, which it is only possible to do with special training. There

is a clear implication that selection for digital material will continue to be circumscribed by practitioners.

#### **7.1.4 Key factors**

Objective four focused on the factors which influence selection, bringing together findings from the literature review, experts and practitioners. The objective was to 'Identify and describe the key factors in selection for digital preservation in memory institutions'. Selection is clearly a very complex process with many influences, and these are summarised in table 17.

It has become clear through this research that personal, social and conceptual factors are key to selection. Relationships that the individual practitioner has are a large influence on their selection activities and especially on their level of engagement with digital material and digital preservation. Positive relationships include those with colleagues where they act as guides to decisions making. Some relationships which potentially undermine the traditional role and responsibility of the practitioner may be difficult and need careful management as this can cause anxiety and a feeling of losing control of a professional activity amongst practitioners. The concept of engagement describes the activities, knowledge and comfort practitioners have with digital material and this research demonstrates this to be an important factor in selection. Barriers to engagement include a lack of time, resources, and management support. There is a need for greater integration of the institution or the institutional goals and workflows with the organisation, along with clear management structures and a senior manager with responsibility for preservation to drive the engagement of both practitioners and the institution forward. These factors, where lacking, are shown to be underlying influences on selection. Structured ways of interacting between departments as well as the development of informal personal contacts would aid selection, preservation and sourcing of material. Practitioners could help themselves by developing such contacts.

#### **7.1.5 Conceptual model of key factors**

The final objective devised to meet the overall aim of this research was to 'construct a conceptual model of key factors and their relationships which influence selection'. Selection has not been conceptualised in this way in the literature and so a conceptual model which describes the key interrelated factors that underpin selection has been developed. Devising this model involved examining key factors and tracing their relationships. The model indicates the complexity and interrelatedness of factors; the material is only one of these. These factors can be seen as mechanisms which produce or affect observable selection outcomes, corresponding with the view from critical realism that mechanisms are the appropriate focus of social research (Danermark et al 2002, p.163). These factors have been discussed in section 6.1 in particular.

## 7.2 Recommendations

1. Clear policy or procedural guidelines and workflows for staff should be devised in institutions to guide the selection and management of digital material.
2. Responsibilities, within institutions and in collaborative activities, for selection need to be made clear so that all stakeholders understand their roles.
3. Senior managers should be aware of their role in championing digital preservation and ensure there are clear lines of communication and management for selection and preservation within the organisation.
4. Clauses identifying selection principles and criteria are included in policies for the benefit of internal and external stakeholders.
5. The need for training to increase familiarity with digital material and digital preservation issues should be recognised and supported by senior management and those organisations which provide training. Whilst there are training opportunities available, such as the DTPT course or those from British Library Preservation Advisory Centre, evidence from this research makes it clear that practitioners are not always able to access them. Therefore the provision of training should be expanded so practitioners in more institutions are able to access it; awareness courses in particular that do not assume familiarity with digital preservation terminology or concepts might be useful. This would increase engagement with digital material and confidence with selection.
6. Although networks of expertise clearly exist, such as the DPC, the DCC, through professional associations or mailing lists, access to these for support with selection and managing digital material is not available to all practitioners, either through lack of awareness of them or through a lack of opportunity to access them. Therefore networks of expertise formal and informal, internal and external, should be maintained and strengthened, with staff being encouraged to participate and support given by management to take part. Practitioners should aim to widen their networks to access expertise elsewhere. Those within the networks should publicise their work to a wider audience and encourage participation.
7. Decisions are made within institutions as to the appropriate skill set for staff and then they received appropriate training. Consciousness of the way expertise will be accessed enables institutional structures to be in place and greater clarity over roles and responsibilities.
8. There should be greater examination of theory in the literature and the theoretical basis for assumptions that are made. There needs to be a greater emphasis in research on the human factors of selection of digital material as well as the technical and organisational aspects.
9. There should be a clarification of terminology, with care being taken in using terms such as appraisal and significant.

10. The concept of a 'lifetime' or recognition of potential time limits on how long material is to be kept for would help institutions to take responsibility for preservation; there should be greater questioning of the assumption of permanent responsibility.

### **7.3 Research contribution**

Selection has not been widely examined in the literature and this study begins to address this deficiency; few other studies have focused solely on selection. Selection has been addressed in the digital preservation literature from a practical viewpoint, but this research has examined conceptual as well as practical factors in selection. The literature review revealed a lack of digital preservation theory that explicitly focused on selection and this research has been a first step in identifying factors that could underpin a theoretical perspective, which would complement the practical guidance given in the literature.

The research has uncovered issues not discussed in the literature. Relationships with other stakeholders, the influence of professionalism and drivers to selection have been explored. Core criteria for selection and core policy clauses which include the content to be selected have been identified. This research has highlighted and begun to investigate the social aspect of selection. It has examined the views of practitioners which were previously neglected. The practitioners added to and clarified many issues of practice which are not addressed in the literature, such as unstated criteria used to select material, and gave a rich picture of the key factors. This research provides an appreciation of the variety and complexity of influences on selection.

This research has developed a basis for conceptualisation of selection in digital preservation. . By developing a conceptual model of selection of key factors, influences that have previously been hidden, such as professionalism and relationships, can be further examined. The model brings together both practical and conceptual factors identified in this research to provide a broad view of selection for digital preservation which is missing from the literature. The model could act as a guide to future selection, enabling all the influential factors to be considered; it is a tool to understand selection and to explain selection decisions. The model communicates the complexity of selection, enables greater understanding of influences and their interrelationships, and can facilitate further discussion and examination of selection.

### **7.4 Suggestions for further work**

1. More needs to be known about the role of the digital curator, how this is developing, whether it is necessary or desirable, and how it could be integrated with other relevant professional roles; these could be investigated further.

2. Other stakeholders were not included in this research due to limitations of time and resources, so further research focusing on their views would complement the views of practitioners examined here. In particular the views of senior managers and technical staff within institutions could be sought, as these have been identified as influential but are unexamined. Also external stakeholders, which as an umbrella term include quite different groups such as users, funders, including the HLF and research councils, local authority councillors and trustees, could usefully be consulted for their views. This would add to the rich picture of selection found in this research.
3. It is apparent from previous research that the criteria used by practitioners and those considered important by users are not always the same. Including users was outside the scope of this study, so examining the views of users on selection, and in particular criteria, in different contexts would provide an interesting counterpoint to the practitioner interviews and allow further refinement of criteria.
4. This research has found that personal and professional networks are a key method for disseminating expertise regarding selection. Further understanding of the structure of professional networks and the methods of disseminating expertise could be of benefit. Questions could for example include is it of greater benefit to provide face to face awareness training or are online materials sufficient? How are networks developing for digital preservation and how could practitioners be encouraged to participate and share expertise?

## **7.5 Reflection and limitations**

The intensive, qualitative approach has inherent limitations that have been discussed in chapter 3. Rather than providing a statistically representative survey of the views of stakeholders, the research has aimed to explore a range of concerns and views from different cultural heritage domains. The research focused on the UK, perhaps limiting its usefulness in other countries with different selection and appraisal traditions. Using critical realism as an underpinning research philosophy focused the researcher on the underlying processes of selection, asking not just what is done but what drives it to be done in certain ways; this has been a particular strength of this study as it has identified a complex picture, including drivers and barriers to selection, unstated criteria and key practical and conceptual factors which underpin selection practice.

The data were gathered through a series of in depth interviews with both managers and curatorial content focused staff. The practitioners were mostly very forthcoming and helpful in their responses. The method of performing the interviews had a strong effect on the quality of the interview. The interviews performed through email were in general less helpful than either the telephone or face to face as it was easy for the respondents to give very short answers and it was difficult for the interviewer to encourage more in depth responses. However some email respondents were very helpful and the interviewer was able to return with further clarifying

questions. The telephone interviews were useful in that sometimes it was simply not possible to meet face to face but some difficulties were experienced as it is very easy to interrupt by accident and it was not possible to encourage people through body language or expressions. The face to face interviews were only difficult when they were made in a public place, such as a cafe, as the background noise occasionally obscured the words of the respondent in the recording.

In spite of these issues, most respondents were very frank and open. The interviewees on the whole were generous with their time, which gave opportunities in the interviews to follow up interesting points providing richer data for the research. The sampling strategy worked effectively and the fact that the researcher had previously made contact with many of the interviewees online through the request for digital preservation policies was particularly effective at gaining their agreement to be interviewed. The lack of museum respondents was a disappointment however. The researcher perhaps could have improved things by separating out the managers from the curators more clearly when sampling, but had to rely to some extent on the description of themselves by the respondents.

## **7.6 Concluding remarks**

The value of this research has been as a systematic investigation of an important yet neglected part of the digital preservation process, through gathering the views of stakeholders that are underrepresented in the literature. Selection as a topic has been under examined and this research has clearly identified the underlying complexity and influences involved in selecting digital material for preservation. Although highly engaged practitioners regard digital as just another type of material, many are clearly anxious and struggling to engage with it. Whilst quantitative research in the form of surveys has been carried out into digital preservation the qualitative approach used here is unusual. The research has gone beyond only asking 'what happens?', although this is important, by asking 'what underlying factors influence what happens?'

## BIBLIOGRAPHY

---

Adcock, E., ed., 1998. *Principles for the care and handling of library material*. Paris: IFLA-PAC.

Ambrose, T., and Paine, C., 1993. *Museum basics*. London: ICOM in conjunction with Routledge.

American Library Association, 2009. *Definition of digital preservation*. [online], Association for Library Collections & Technical Services. Available at: <http://www.ala.org/alcts/resources/preserv/2009def>, [accessed 26.04.13].

Anderson B., 1999. What's worth digitizing? *Behavioral & Social Sciences Librarian*, 18(1), pp.75-77.

Anderson, D., 2011. *A layman's guide to the KEEP legal studies*. [online]. Available at: <http://www.keep-project.eu/ezpub2/index.php?/eng/About-KEEP/News/Layman-s-Guide-to-Legal-Studies-Published>, [accessed 24.05.13].

Anderson, D., Delve, J., and Pinchbeck, D., 2010. Toward a workable emulation-based preservation strategy: rationale and technical Metadata. *New review of Information Networking* [online], 15(2), pp. 110-131. Available at: <http://dx.doi.org/10.1080/13614576.2010.530132> [accessed 30.04.13].

Anderson, R., ed. 2006. *The e-resources management handbook*. [e-book] UKSG. Available at: <http://uksq.metapress.com/content/5vgwxpbtr393fhl/?p=a3b8368715754384b321afe354bb1706&pi=26>, [accessed 24.05.13].

Angevaare, I., Gallinger, M., Anderson, M., Giaretta, D., and Halbert, M., 2012. Organizational alignment. In: McGovern, N.Y. & Skinner, K., eds. *Aligning national approaches to digital preservation*. Atlanta, Georgia: Educopia Institute Publications, pp.89-114.

Archaeology Data Service, [n.d.]. *About us*. <http://archaeologydataservice.ac.uk/about/contact>, [accessed 30.04.13].

Archives and Records Association, 2012. *Code of conduct*. [online]. Available at: <http://www.archives.org.uk/membership/code-of-conduct.html> [accessed 30.04.13].

*Archives for the 21<sup>st</sup> century*. 2009. Department of Justice, (Cmnd.7744) London: HMSO.

Arthur, J., 2009. *Survey on preservation training requirements*. [online], London: British Library. Available at: <[www.bl.uk/blpac/pdf/survey.pdf](http://www.bl.uk/blpac/pdf/survey.pdf)>, [accessed 24.05.13].

Arts Council England, [n.d.]. *Who we are*.< <http://www.artscouncil.org.uk/who-we-are/>>, [accessed 30.04.13].

Arts Council England, 2012. *Designated collections list*. [online], Arts Council England. Available at: [http://www.artscouncil.org.uk/media/uploads/pdf/Designation\\_List.pdf](http://www.artscouncil.org.uk/media/uploads/pdf/Designation_List.pdf), [accessed 26.04.13].

Asheim, L., 1953. Selection and censorship: a reappraisal. *Wilson Library Bulletin*, [online], 28. Available at: <<http://www.ala.org/offices/oif/basics/notcensorship>>, [accessed 26.04.13].

Assman, J., 2008. Communicative and cultural memory. In: Erll, A., & Nunning, A., eds. *Cultural memory studies: an international and interdisciplinary handbook*. Berlin: Walter de Gruyter, pp. 109-118.

Atkinson, R., 1986. Selection for preservation: a materialistic approach. *Library Resources & Technical Services*, 30, pp.341-53.

Atkinson, R., 1993. Access, ownership, and the future of collection development. In: Johnson, P., & MacEwan, B., eds. *Collection management and development: issues in an electronic age*. Chicago: American Library Association, pp. 92-109.

Austin, T. and Richards, J., 2011. *Archaeology data service: preservation policy*. [online], Archaeology Data Service. Available at: <<http://archaeologydataservice.ac.uk/attach/preservation/PreservationPolicyV1.3.1.pdf>>, [accessed 26.04.13].

Ayre, C., and Muir, A., 2004. *Right to preserve? The copyright and licensing for digital preservation project final report*. Loughborough: Loughborough University.

Ayris, P., 1998. Guidance for selecting materials for digitisation. *Joint RLG and NPO Preservation Conference: Guidelines for Digital Imaging*. [online], Warwick. Available at: [http://eprints.ucl.ac.uk/492/1/paul\\_ayris3.pdf](http://eprints.ucl.ac.uk/492/1/paul_ayris3.pdf) [accessed 26.04.13].

Ayris, P., Davies, R., McLeod, R., Miao, R., Shenton, H., and Wheatley, P., 2008. *The Life<sup>2</sup> final project report*. [online], London: LIFE Project. Available at: <<http://eprints.ucl.ac.uk/11758/>>, [accessed 30.04.13].

Bailey, S., 2008. *Managing the crowd: rethinking records management for the Web 2.0 world*. London: Facet.

Barateiro, J., Antunes, G., Freitas, F., and Borbinha, J., 2010. Designing digital preservation solutions: a risk-management-based approach. *The International Journal of Digital Curation* [online], 5(1), pp. 4-17. Available at: <<http://ijdc.net/index.php/ijdc/article/view/143>>, [accessed 30.04.13].

Barwick, J., Dearnley, J., and Muir, A., 2008. *The barriers to the preservation of digital games: questions on cultural significance*. [online], DOCAM and Media@McGill. Available at: <https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/4988>, [accessed 30.04.13].

BBC, 2006. *WW2 people's war*. <http://www.bbc.co.uk/ww2peopleswar/> [accessed 26.04.13].

Beagrie, C. Ltd., 2012. KRDS/I2S2 digital preservation benefit analysis tools. [online]. Available at: <<http://beagrie.com/krds-i2s2.php>> [accessed 13.05.13].

Beagrie, N., and Greenstein, D., 1998. *A strategic policy framework for creating and preserving digital collections*. [online], London: British Library Research and Innovation Centre. Available at: <http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/framework.pdf> [accessed 26.04.13].

Beagrie, N., 2006. Digital curation for science, digital libraries, and individuals. *The International Journal of Digital Curation* [online], 1(1), pp.3-16. Available at: <<http://www.ijdc.net/index.php/ijdc/article/view/6>>, [accessed 30.04.13].

Beagrie, N., Rettberg, N., and Williams, P., 2008. Digital preservation policy: a subject of no importance? In: British Library. *The Fifth International Conference on Preservation of Digital Objects Joined Up and Working: Tools and Methods for Digital Preservation*. London, UK, 29-30 September 2008. London: British Library.

Beagrie, N., Semple, N., Williams, P., and Wright, R., 2008. *Digital preservation policies study part 1*. [online], JISC. Available at: [http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy\\_p1finalreport.pdf](http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf) [accessed 26.04.13].

Beagrie, N., Lavoie, B., and Woollard, M., 2010. *KRDS2 Data survey - selection criteria*. [online], JISC. Available at: <http://www.jisc.ac.uk/media/documents/publications/reports/2010/keepingresearchdatasafe2.pdf> [accessed 26.04.13].

Bearman, D. and Trant, J., 1998. Economic, social, technical models for digital libraries of primary resources: The example of the art museum image consortium (AMICO). *New Review of Information Networking* [online], 4(1), pp. 71-91. Available at: <http://www.tandfonline.com/doi/abs/10.1080/13614579809516918#.UZCx-cpu18F> [accessed 13.05.13].

Bearman D., 2007. Addressing selection and digital preservation as systemic problems. In: de Lusenet, Y., & Wintermans, V., eds. *Preserving the digital heritage principles and policies*. Amsterdam: Netherlands National Commission for UNESCO, pp. 26-44.

Beedham, H., Missen, J., Palmer, M., and Ruusalepp, R. 2005. *Assessment of UKDA and TNA compliance with OAIS and METS standards* [online], UK Data Archive. Available at: <http://www.jisc.ac.uk/media/documents/programmes/preservation/oaismets.pdf> [accessed 13.05.13].

Berger, S., 2009. The evolving ethics of preservation: redefining practices and responsibilities in the 21st century. *The Serials Librarian*, 57(1), pp. 57-68.

Besek, J.M., LeFurgy, W.G., Weston, C.D., Muir, A., Atkinson, B., Carroll, E., Coates, J., Fitzgerald, B., 2008. *International study on the impact of copyright law on digital preservation* [online], Available at: [http://www.digitalpreservation.gov/documents/digital\\_preservation\\_final\\_report2008.pdf](http://www.digitalpreservation.gov/documents/digital_preservation_final_report2008.pdf) [accessed 26.04.13].

Bhaskar, R., 1975. *Realist theory of science*. York: Alma.

Bibliothèque nationale de France, 2011. *Digital legal deposit: four questions about web archiving at the BnF*. [http://www.bnf.fr/en/professionals/digital\\_legal\\_deposit/a.digital\\_legal\\_deposit\\_web\\_archiving.html](http://www.bnf.fr/en/professionals/digital_legal_deposit/a.digital_legal_deposit_web_archiving.html) [accessed 26.04.13].

Billenness, C., 2011. *The future of the past – shaping new visions for EU-research in digital preservation*, European Commission, Luxembourg, 4-5 May 2011. [online]. Available at: [http://cordis.europa.eu/fp7/ict/telearn-digicult/future-of-the-past\\_en.pdf](http://cordis.europa.eu/fp7/ict/telearn-digicult/future-of-the-past_en.pdf) [accessed 13.05.13].

Birrell, D., Dobрева, M., Dunsire, G., Griffiths, J.R., Hartley, R.J., and Menzies, K., 2011. The DISCmap project digitisation of special collections: mapping assessment, prioritisation. *New Library World*, 112(1/2), pp.19-44.

- Black, A., 2006. The people's university: models of public library history. *In*: Black, A. & Hoare, P., eds. *The Cambridge history of libraries in Britain and Ireland Vol.III 1850-2000*, pp. 24-39.
- Black, A. and Hoare, P., eds., 2006. *The Cambridge history of libraries in Britain and Ireland Vol.III 1850-2000*. Cambridge: Cambridge University Press.
- Blaikie, N., 1993. *Approaches to social enquiry*. Cambridge: Polity Press.
- Blaikie, N., 2000. *Designing social research: the logic of anticipation*. Cambridge: Polity Press.
- Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010. *Sustainable economics for a digital planet: ensuring long term access to digital information*. [online]. Available at: [http://brtf.sdsc.edu/biblio/BRTF\\_Final\\_Report.pdf](http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf), [accessed 26.04.13].
- Blumer, H., 1954. What is wrong with social theory? *American Sociological Review* 19(1), pp.3-10.
- Boylan, P., ed., 2004. *Running a museum: a practical handbook*. Paris: ICOM.
- Breitbach, W. and Lambert, J.E., 2011. Patron-driven ebook acquisition. *Computers in Libraries*, 31(6), pp. 16-20.
- British Library, [n.d.]. *Legal deposit*. <http://www.bl.uk/aboutus/stratpolprog/legaldep/>, [accessed 26.04.13].
- Brown, A., 2009. *A digital preservation policy for parliament*. [online], Parliamentary Archives. Available at: <http://www.parliament.uk/documents/upload/digitalpreservationpolicy1.0.pdf>, [accessed 26.04.13].
- Bryman, A., 2004. *Social research methods*. 2nd ed. Oxford: Oxford University Press.
- Bryman, A., 2008. *Social research methods*. 3rd ed. Oxford: Oxford University Press.
- Bryman, A., 2012. *Social research methods*. 4th ed. Oxford: Oxford University Press.
- Cameron, F., 2007. Beyond the cult of the replicant – museums and historical digital objects: traditional concerns, new discourses. *In*: Cameron, F. and Kenderdine, S., eds., *Theorizing digital cultural heritage*. Cambridge, Mass.: The MIT Press, pp. 49-76.

Cameron, F. and Kenderdine, S., eds., 2007. *Theorizing digital cultural heritage*. Cambridge, Mass.: The MIT Press.

Cameron, F., 2008. The politics of heritage authorship: the case of digital heritage collections. In: Kalay, Y.E., Kvan, T. And Affleck, J., eds. *New heritage*. London: Routledge. pp.170-184.

Canadian Heritage Information Network, 2006. *Concepts for developing digital preservation policies*. <[http://www.pro.rcip-chin.gc.ca/sommaire-summary/politiques\\_preservation-preservation\\_policies-eng.jsp](http://www.pro.rcip-chin.gc.ca/sommaire-summary/politiques_preservation-preservation_policies-eng.jsp)>, [accessed 01.05.13].

Candela, L., Castelli, D., Ferro, N., Ioannidis, Y., Koutrika, G., Meghini, C., Pagano, P., Ross, S., Soergel, D., Agosti, M., Dobрева, M., Katifori, V., and Schuldt, H., 2007. *The DELOS digital library reference model - foundations for digital libraries. Version 0.98*. [online]. Available at: [http://www.delos.info/index.php?option=com\\_content&task=view&id=345#reference\\_model](http://www.delos.info/index.php?option=com_content&task=view&id=345#reference_model), [accessed 24.05.13].

Cardy, L., and Rumsey, S., 2006. Purchasing models for scholarly materials in the UK: a case-study. In: Anderson, R., ed. *The e-resources management handbook*. [e-book] UKSG. pp. 67-77.

CCSDS, 2004. *Recommended practice for a producer-archive interface methodology abstract standard*. [online]. Available at: < [public.ccsds.org/publications/archive/651x0m1.pdf](http://public.ccsds.org/publications/archive/651x0m1.pdf)>, [accessed 24.05.13].

CCSDS, 2012. *Reference model for an open archival information system (OAIS)*. [online]. Available at: < [public.ccsds.org/publications/archive/650x0m2.pdf](http://public.ccsds.org/publications/archive/650x0m2.pdf)>, [accessed 24.05.13].

Cedars Project, 2002. *Cedars guide to digital collection management*. [online], The Cedars Project. Available at: <http://www.webarchive.org.uk/wayback/archive/20050410120000/http://www.leeds.ac.uk/cedars/guideto/collmanagement/guidetocolman.pdf>, [accessed 01.05.13].

Charmaz, K., 2006. *Constructing grounded theory: a practical guide through qualitative analysis*. London: Sage.

Checkley-Scott, C. And Thompson, D., 2007. *Wellcome library preservation policy for materials held in collections*. [online], Wellcome Library. <<http://library.wellcome.ac.uk/assets/wtx038065.pdf>>, [accessed 01.05.13].

Chowdhury, G., 2010. From digital libraries to digital preservation research: the importance of users and context. *Journal of Documentation*, 66(2), pp.207-223.

CILIP, 2012a. *Ethical principles for library and information professionals*. <<http://www.cilip.org.uk/get-involved/policy/ethics/pages/principles.aspx>>, [accessed 24.05.13].

CILIP, 2012b. *Code of professional practice for library and information professionals*. <<http://www.cilip.org.uk/get-involved/policy/ethics/pages/code.aspx>>, [accessed 24.05.13].

CILIP, 2012c. *Professional knowledge and skills base*. [online], London: CILIP. Available at: <<http://www.cilip.org.uk/jobs-careers/professional-knowledge-and-skills-base/pages/professional%20knowledge%20and%20skills%20base.aspx>>, [accessed 24.05.13].

Clayton, P.R. and Gorman, G. E., 2001. *Managing information resources in libraries: collection management in theory and practice*. London: Library Association.

Clifton, G., 2005. Risk and the preservation management of digital collections. *International Preservation News*. [online]. 36. pp.21-23. Available at: <[archive.ifla.org/VI/4/news/ipnn36.pdf](http://archive.ifla.org/VI/4/news/ipnn36.pdf)>, [accessed 24.05.13].

Collections Trust, 2011. *SPECTRUM 4.0*. [online], London: Collections Trust. Available at: <<http://www.collectionslink.org.uk/spectrum-standard>>, [accessed 24.05.13].

Conway, P., 1996. *Preservation in the digital world*. [online], Washington, DC: Council on Library Resources. Available at: <http://www.clir.org/pubs/reports/conway2/reports/conway2/index.html> (accessed 01.05.13).

Conway, P., 2010. Preservation in the age of Google: digitization, digital preservation, and dilemmas. *Library Quarterly*, 80(1), pp. 61–79.

Cook, T., 1997. What is past is prologue: a history of archival ideas since 1898, and the future paradigm shift. *Archivaria*, 43, pp.17-63.

Cook, T., 2011. We are what we keep; we keep what we are': archival appraisal past, present and future. *Journal of the Society of Archivists*, 32(2), pp.173-189.

Craig, B., 2004. *Archival appraisal: theory and practice*. Munchen: K.G.Saur.

- Creswell, J.W., 2009. *Research design: qualitative, quantitative, and mixed methods approaches*. 3rd ed. Thousand Oaks, CA.; London: Sage Publications.
- Currall, J., and McKinney, P., 2006. Investing in value a perspective on digital preservation. *D-Lib Magazine*. [online]. 12(4). Available at: <<http://www.dlib.org/dlib/april06/mckinney/04mckinney.html>>, [accessed 24.05.13].
- Currall, J., McKinney, P., and Keenan, J., 2007. *Final report*. [online]. Available at: <[www.jisc.ac.uk/media/documents/programmes/preservation/espidaprojectjiscfinalreport.pdf](http://www.jisc.ac.uk/media/documents/programmes/preservation/espidaprojectjiscfinalreport.pdf)>, [accessed 24.05.13].
- Danermark, B., Ekstrom, M., Jakobsen, L., and Karlsson, J.C., 2002. *Explaining society: critical realism in the social sciences*. London: Routledge.
- Dappert A., Ballaux, B., Mayr, M., and Van Bussel, S., 2008. *Report on policy and strategy models for libraries, archives and data centres*. [online]. Available at: [http://www.planets-project.eu/docs/reports/Planets\\_PP2\\_D2\\_ReportOnPolicyAndStrategyModelsM24\\_Ext.pdf](http://www.planets-project.eu/docs/reports/Planets_PP2_D2_ReportOnPolicyAndStrategyModelsM24_Ext.pdf), (accessed 01.05.13).
- Dappert, A., and Farquhar, A., 2009. *Significance is in the eye of the stakeholder*. [online]. Available at: <[www.planets-project.eu/docs/papers/Dappert\\_Significant\\_Characteristics\\_ECDL2009.pdf](http://www.planets-project.eu/docs/papers/Dappert_Significant_Characteristics_ECDL2009.pdf)>, [accessed 24.05.13].
- Day, M., 2003. *Collecting and preserving the World Wide Web*. [online]. UKOLN. Available at: <[www.jisc.ac.uk/uploaded\\_documents/archiving\\_feasibility.pdf](http://www.jisc.ac.uk/uploaded_documents/archiving_feasibility.pdf)>, [accessed 24.05.13].
- Day, M., 2005. *Metadata*. [online], DCC. Available at: <<http://www.dcc.ac.uk/resources/curation-reference-manual/completed-chapters/metadata>>, [accessed 01.05.13].
- Day, M., Pennock, M., and Allinson, J., 2007. Co-operation for digital preservation and curation: collaboration for collection development in institutional repository networks. DigCCur2007 International Symposium in Digital Curation, Chapel Hill, NC, USA, 18-20 April 2007. Available at: [http://www.ils.unc.edu/digccurr2007/papers/dayPennock\\_paper\\_9-3.pdf](http://www.ils.unc.edu/digccurr2007/papers/dayPennock_paper_9-3.pdf) [accessed 01.05.13].
- Deegan, M., and Tanner, S., 2006. *Digital preservation*. London: Facet.

Del Pozo, N., Stawowczyk Long, A., and Pearson, D., 2010. "Land of the lost": a discussion of what can be preserved through digital preservation. *Library Hi Tech*, 28(2), pp.290 - 300.

De Lusenet, Y. and Wintermans, V., eds., 2005. *Preserving the digital heritage principles and policies*. Amsterdam: Netherlands National Commission for UNESCO.

Dempsey, L., 1999. Scientific, industrial, and cultural heritage: a shared approach. A research framework for digital libraries, museums and archives. *Ariadne* [online], 22. Available at: <<http://www.ariadne.ac.uk/issue22/dempsey/>>, [accessed 01.05.13].

Dempsey, L., 2006. The (digital) library environment: ten years after. *Ariadne* [online], 46. Available at: <<http://www.ariadne.ac.uk/issue46/dempsey/>>, [accessed 01.05.13].

Denzin, K.N. and Lincoln, Y.S., 2000. *The handbook of qualitative research*. 2nd ed. London: Sage Publications.

Department of Culture Arts and Leisure, n.d. *About us*. <[http://www.dcalni.gov.uk/about\\_us.htm](http://www.dcalni.gov.uk/about_us.htm)>, [accessed 24.05.13].

Department for Culture Media and Sport, n.d.a. *Arts and culture*. <<https://www.gov.uk/government/topics/arts-and-culture>>, [accessed 24.05.13].

Department for Culture Media and Sport, n.d.b. *English Heritage*. <<https://www.gov.uk/government/organisations/english-heritage>>, [accessed 24.05.13].

Department for Culture Media and Sport, n.d.c. *The British Library*. <<https://www.gov.uk/government/organisations/british-library>>, [accessed 24.05.13].

Department for Culture Media and Sport, 2010. *Improving efficiency and transparency: DCMS cuts quangos*. Available at: <<https://www.gov.uk/government/news/improving-efficiency-and-transparency-dcms-cuts-quangos>>, [accessed 01.05.13].

Department for Culture Media and Sport, 2011. *Government response to the public consultation on the draft regulations and guidance for non-print legal deposit*. [online], Available at: <http://www.culture.gov.uk/publications/8029.aspx> (accessed 01.05.13).

Det Kongelige Bibliotek, 2010. *Policy for the long term preservation of digital materials at the Royal Library*. [online], DetKongeligeBibliotek. Available at: [http://www.kb.dk/export/sites/kb\\_dk/da/kb/downloadfiler/preservationpolicy\\_kb\\_2009.pdf](http://www.kb.dk/export/sites/kb_dk/da/kb/downloadfiler/preservationpolicy_kb_2009.pdf), [accessed 01.05.13].

*Digital Britain: Final Report*, 2009. Department for Culture, Media and Sport and Department for Business Innovations and Skills, (Cmnd. 7650) London: HMSO.

Digital Curation Centre, 2009. *DCC DIFFUSE standards frameworks*.  
<<http://www.dcc.ac.uk/resources/standards/diffuse>>, [accessed 24.05.13].

Digital Curation Centre, 2010. *What is digital curation?*  
<http://www.dcc.ac.uk/digital-curation/what-digital-curation>, [accessed 01.05.13].

Digital Preservation Coalition, 2004. *Training needs analysis final report*. [online] York: Digital Preservation Coalition. Available at:<[www.jisc.ac.uk/media/documents/programmes/preservation/trainingneedsfinalreport.pdf](http://www.jisc.ac.uk/media/documents/programmes/preservation/trainingneedsfinalreport.pdf)>, [accessed 24.05.13].

Digital Preservation Coalition, 2006. *Decision tree for selection of digital materials for long-term retention*. [online], York: Digital Preservation Coalition. Available at:<http://www.dpconline.org/advice/preservationhandbook/decision-tree> [accessed 01.05.13].

Digital Preservation Coalition, 2008. *Digital preservation handbook*. [online], York: Digital Preservation Coalition. Available at: <http://www.dpconline.org/advice/preservationhandbook> [accessed 01.05.13].

DTPT, n.d. *About*. <<http://www.dptp.org/about-2/>>, [accessed 24.05.13].

Donnelly, M., Innocenti, P., McHugh, A., and Ruusalepp, R., 2009. *DRAMBORA interactive: user guide*. [online], DCC. Available at: <<http://www.era.lib.ed.ac.uk/handle/1842/3328>>, [accessed 24.05.13].

Doom, V., 2004. Selection criteria. In: International Council on Archives. *Draft manual on appraisal*. [online]. Available at: <<http://www.ica.org/11274/toolkits-guides-manuals-and-guidelines/draft-manual-on-appraisal.html>>, [accessed 24.05.13].

DSpace@Cambridge, 2008. *Policies*.  
<http://www.lib.cam.ac.uk/repository/about/policies.html#preservation>>, [accessed 01.05.13].

Duranti, L., 1994. The concept of appraisal and archival theory. *American Archivist*, 57, pp.328 - 344.

Duranti, L. and Preston, R., 2008. *International research on permanent authentic records in*

*electronic systems (INTERPARES) 2: experiential, interactive and dynamic records*. [online], Padova, Italy: Associazione Nazionale Archivistica Italiana. Available at: <<http://www.interpares.org/ip2/book.cfm>>, [accessed 24.05.13].

Durrant, S., 2008. *Long-term preservation results from a survey investigating preservation strategies amongst ALPSP publisher members*. [online], Association of Learned and Professional Society Publishers. Available at: <http://www.alpsp.org/Ebusiness/ProductCatalog/longtermpreservation.aspx?ID=39> [accessed 01.05.13].

Eastwood, T., 2004. Appraising digital records for long-term preservation. *Data Science Journal*, [online], 3. Available at: [http://www.interpares.org/display\\_file.cfm?doc=ip1\\_dissemination\\_jar\\_eastwood\\_data\\_science\\_journal\\_3\\_2004.pdf](http://www.interpares.org/display_file.cfm?doc=ip1_dissemination_jar_eastwood_data_science_journal_3_2004.pdf)>, [accessed 01.05.13].

Edwards, A., Matthews, G., and Nankivell, C., 2000. *Developing a national strategy for preservation surrogates*. Library and Information Commission Research Report 54. London: Library and Information Commission.

Eden, P., 1997. Concern for the future: preservation management in libraries and archives. *Journal of Librarianship and Information Science*, 29(3), pp.121-129.

Emmott, S., 2008. Preservation of web resources: making a start. *Ariadne*. [online]. 56. Available at: <http://www.ariadne.ac.uk/issue56/jisc-powr-rpt>, [accessed 24.05.13].

*Encyclopedia of Library and Information Sciences*. 3rd ed. [online]. New York: Taylor and Francis. Available at: <<http://www.tandfonline.com/doi/book/10.1081/E-ELIS3?doi=10.1081/E-ELIS3&sort=contrib>>, [accessed 24.05.13].

ERPANET, 2003. *Digital preservation policy tool*. [online], ERPANET. Available at: <<http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf>>, [accessed 01.05.13].

Erll, A. and Nunning, A., 2008. *Cultural memory studies: an international and interdisciplinary handbook*. Berlin: Walter de Gruyter.

Farrell, S., ed., 2010. *A guide to web preservation*. [online], JISC. Available at: <[jiscpowr.jiscinvolve.org/files/2010/06/Guide-2010-final.pdf](http://jiscpowr.jiscinvolve.org/files/2010/06/Guide-2010-final.pdf)>, [accessed 24.05.13].

Feather, J., 1996. *Preservation and the management of library collections*. London: Library Association Publishing.

Feather, J., 2004. *Managing preservation for libraries and archives: current practice and future developments*. Aldershot: Ashgate.

Feather, J., 2006. Managing the documentary heritage: issues for the present and future. *In*: Gorman, G.E., & Shep, S.J., eds. *Preservation management*. London: Facet Publishing, pp.1-18.

Feeney, M. and National Preservation Office, 1999. *Digital culture: maximising the nation's investment: a synthesis of JISC/NPO studies on the preservation of electronic materials*. London: National Preservation Offices, British Library.

Ferreira, M., Baptista, A.A., and Ramalho, J.C., 2006. A foundation for automatic digital preservation. *Ariadne*. [online]. 48. Available at: <<http://www.ariadne.ac.uk/issue48/ferreira-et-al>>, [accessed 24.05.13].

Flow Associates, and The Collections Link, 2010. *Mapping the use of digital technologies in the heritage sector*. [online], HLF. Available at: [www.hlf.org.uk/aboutus/howwework/Documents/HLF\\_digital\\_review.pdf](http://www.hlf.org.uk/aboutus/howwework/Documents/HLF_digital_review.pdf)>, [accessed 24.05.13].

Flynn, S.J.A., 2001. The records continuum model in context and its implications for archival practice. *Journal of the Society of Archivists*, 22(1), pp.79-93.

Foot, M.M., 2001. *Building blocks for a preservation policy*. [online], London: National Preservation Office. Available at: <<http://www.bl.uk/blpac/pdf/blocks.pdf>>, [accessed 01.05.13].

Fresco, M., 1996. *Long term preservation of electronic materials: a JISC/British Library workshop as part of the electronic libraries programme (eLib) organised by UKOLN, 27th and 28th November 1995 at the University of Warwick*. British Library Research and Development Report, 6238. [online], British Library Research and Development Department. Available at: <<http://www.ukoln.ac.uk/services/papers/bl/rdr6238/>>, [accessed 01.05.13].

Gantz, J., and Reinsel, D., 2011. *Extracting value from chaos*. [online], IDC. Available at: <[uk.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf](http://uk.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf)>, [accessed 24.05.13].

Giaretta, D., 2011. *Advanced digital preservation*. Heidelberg: Springer.

Gibby, R. and Green, A., 2008. Electronic legal deposit in the United Kingdom. *New Review of Academic Librarianship* [online], 14(1), pp.45-70. Available at:

<http://www.informaworld.com/smpp/content~db=all?content=10.1080/13614530802519022>, [accessed 01.05.13].

Gibson, L. and Pendlebury, J.R., eds., 2009. *Valuing historic environments*. Farnham: Ashgate.

Gibson, L., 2009. Cultural landscapes and identity. In: Gibson, L. & Pendlebury, J.R., eds. *Valuing historic environments*. Farnham: Ashgate, pp.67-91.

Given, L.M. and McTavish, L., 2010. What's old is new again: the reconvergence of libraries, archives, and museums in the digital age. *The Library Quarterly*, 80(1), pp. 7-32.

Gladney, H.M., 2008. *Durable digital objects rather than digital preservation*. [online]. Available at: < eprints.erpanet.org/146/01/Durable.pdf>, [accessed 24.05.13].

Glaser, B.G. and Strauss, A.L., 1967. *The discovery of grounded theory: strategies for qualitative research*. New York: Aldine de Gruyter.

Gorman, G.E. and Shep, S.J., 2006. *Preservation management for libraries, archives and museums*. London: Facet Publishing.

Gorman, M., 2007. The wrong path and the right path. The role of libraries in access to, and preservation of, cultural heritage. *New Library World*, 108(11/12), pp. 479-489.

Gould, S. and Ebdon, R., 1999. *IFLA/UNESCO Survey on digitisation and preservation*. [online], <[http://www.unesco.org/webworld/mdm/survey\\_index\\_en.html](http://www.unesco.org/webworld/mdm/survey_index_en.html)>, [accessed 01.05.13].

Gov.uk, n.d. *Departments, agencies and public bodies*. <<https://www.gov.uk/government/organisations#department-for-culture-media-sport>>, [accessed 24.05.13].

Gowers, A., 2006. *Gowers review of intellectual property*. [online], HM Treasury. Available at: <[http://webarchive.nationalarchives.gov.uk/+http://www.hm-treasury.gov.uk/d/pbr06\\_gowers\\_report\\_755.pdf](http://webarchive.nationalarchives.gov.uk/+http://www.hm-treasury.gov.uk/d/pbr06_gowers_report_755.pdf)>, [accessed 01.05.13].

Grace, S., Knight, G., and Montague, L., 2009. *Final Report*. [online], JISC. Available at: <[www.significantproperties.org/inspect-finalreport.pdf](http://www.significantproperties.org/inspect-finalreport.pdf)>, [accessed 24.05.13].

Gracy, K.F., 2007. Moving image preservation and cultural capital. *Library Trends*, 56(1), pp. 183-197.

Graeber, D., 2001. *Toward an anthropological theory of value: the false coin of our own dreams*. New York: Palgrave Macmillan.

Graham, B. and Howard, P., eds., 2008. *The Ashgate research companion to heritage and identity*. Aldershot: Ashgate.

Gray, D.E., 2009. *Doing research in the real world*. Los Angeles, Calif.; London: Sage.

Great Britain, 1988. *Copyright, Designs and Patent Act 1988*. London: HMSO. Available at: [http://www.opsi.gov.uk/acts/acts1988/ukpga\\_19880048\\_en\\_1.htm](http://www.opsi.gov.uk/acts/acts1988/ukpga_19880048_en_1.htm), [accessed 01.05.13].

Great Britain, 1958. *Public Records Act 1958*. London: HMSO. Available at: <<http://www.legislation.gov.uk/ukpga/Eliz2/6-7/51>>, [accessed 24.05.13].

Great Britain, 1998. *Data Protection Act 1998 (Chapter 29)*. London: HMSO. Available at: <[http://www.opsi.gov.uk/acts/acts1998/ukpga\\_19980029\\_en\\_1](http://www.opsi.gov.uk/acts/acts1998/ukpga_19980029_en_1)>, [accessed 01.05.13].

Great Britain, 2000. *Freedom of information Act 2000 (Chapter 36)*. London: HMSO. Available at: <[http://www.opsi.gov.uk/acts/acts2000/ukpga\\_20000036\\_en\\_1](http://www.opsi.gov.uk/acts/acts2000/ukpga_20000036_en_1)>, [accessed 01.05.13].

Great Britain, 2003. *Legal deposit Libraries Act 2003 (Chapter 28)*. London: HMSO. Available at: <<http://www.legislation.gov.uk/ukpga/2003/28/contents>>, [accessed 01.05.13].

Great Britain, 2010. *Digital Economy Act 2010 (Chapter 24)*. London: HMSO. Available at: <[http://www.opsi.gov.uk/acts/acts2010/ukpga\\_20100024\\_en\\_1](http://www.opsi.gov.uk/acts/acts2010/ukpga_20100024_en_1)>, [accessed 01.05.13].

Hall, S., 2005. Whose heritage? Un-settling 'the heritage', re-imagining the post-nation. *In*: Littler, J. and Naidoo, R., eds. *The politics of heritage: the legacies of 'race'*. London: Routledge, pp. 23-35.

Hampshire Records Office, 2011. *Digital preservation policy*. <<http://www3.hants.gov.uk/archives/hro-policies/hro-digital-preservation-policy.htm>>, [accessed 01.05.13].

Hargreaves, I., 2011. *Digital opportunity. A review of intellectual property and growth*. [online], Available at: <http://www.ipo.gov.uk/ipreview-finalreport.pdf>, [accessed 01.05.13].

Harré, R., 1970. *The Principles of Scientific Thinking*. London: Macmillan.

- Harrison, M., 2005. *Accession records*. [online], Collections Link. Available at: <<http://www.collectionslink.org.uk/programmes/spectrum/594-accession-records>>, [accessed 01.05.13].
- Harvey, D.R., 2005. *Preserving digital materials*. München: K.G. Saur.
- Harvey, R., 2007. *DCC digital curation manual instalment on appraisal and selection*. [online], HATII, University of Glasgow; University of Edinburgh; UKOLN, University of Bath; Council for the Central Laboratory of the Research Councils. Available at: [http://www.dcc.ac.uk/webfm\\_send/121](http://www.dcc.ac.uk/webfm_send/121), [accessed 01.05.13].
- Harvey, R. and Thompson, D., 2010. Automating the appraisal of digital materials. *Library Hi Tech* [online], 28(2), pp.313-322. Available at: <[10.1108/07378831011047703](https://doi.org/10.1108/07378831011047703)>, [accessed 01.05.13].
- Haynes, D., Streatfield, D., Jowett, T., and Blake, M., 1997. *Responsibility for digital archiving and long term access to digital data*. JISC/NPO Studies on the Preservation of Electronic Materials. British Library Research and Innovation Report, 67. [online], British Library Research and Innovation Centre. Available at: <<http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/digital-preservation.pdf>>, [accessed 01.05.13].
- Hazen, D., Horrell, J., and Merrill-Oldham, J., 1998. *Selecting research collections for digitization*. [online], Washington DC: Council on Library and Information Resources. Available at: <http://www.clir.org/pubs/reports/hazen/pub74.html>, [accessed 01.05.13].
- Hedstrom, M., 2002. Archives, memory and interfaces with the past. *Archival Science*, 2(1-2), pp.21-43.
- HLF, 2013a. *What we fund*. <<http://www.hlf.org.uk/HowToApply/whatwefund/Pages/whatwefund.aspx>>, [accessed 24.05.13].
- HLF, 2013b. *About us*. <<http://www.hlf.org.uk/aboutus/Pages/AboutUs.aspx>>, [accessed 2.05.13].
- Heslop, H., Davis, S., and Wilson, A., 2002. *An approach to the preservation of digital records*. [online], National Archives of Australia. Available at: <[www.naa.gov.au/Images/An-approach-Green-Paper\\_tcm16-47161.pdf](http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm16-47161.pdf)>, [accessed 24.05.13].

- Hewison, D. and Holden, J., 2004. *Challenge and change: HLF and cultural value*. [online], HLF. Available at: <[www.hlf.org.uk/aboutus/howwework/Documents/ChallengeandChange\\_CulturalValue.pdf](http://www.hlf.org.uk/aboutus/howwework/Documents/ChallengeandChange_CulturalValue.pdf)>, [accessed 24.05.13].
- Higgins, S., 2008. The DCC curation lifecycle model. *The International Journal of Digital Curation* [online], 3(1). Available at: <<http://www.ijdc.net/index.php/ijdc/article/viewFile/69/48>>, [accessed 01.05.13].
- Hockx-Yu, H. and Knight, G., 2008. What to preserve? Significant properties of digital objects. *The International Journal of Digital Curation* [online], 3(1). Available at: <<http://www.ijdc.net/index.php/ijdc/article/viewFile/70/49>>, [accessed 01.05.13].
- Holden, J., 2004. *Capturing cultural value*. [online], Demos. Available at: <<http://www.demos.co.uk/files/CapturingCulturalValue.pdf?1240939425>>, [accessed 01.05.13].
- Hole, B., Wheatley, P., Lin, L., McCann, P. and Aitken, B., 2010. The Life3 predictive costing tool for digital collections. *New Review of Information Networking* [online]. 15(2). pp. 81-93. Available at: <<http://dx.doi.org/10.1080/13614576.2010.526014>>, [accessed 24.05.13].
- Hudson, C., 2012. The digital museum. In: Hughes L.M., ed. *Evaluating and measuring the value, use and impact of digital collections*. London: Facet. pp.35-48.
- Hughes L.M., ed., 2012. *Evaluating and measuring the value, use and impact of digital collections*. London: Facet.
- Hutter, M. and Rizzo, I., eds., 1997. *Economic perspectives on cultural heritage*. Basingstoke: Macmillan.
- Hutter, M. and Throsby, M., 2008. Value and valuation in art and culture: introduction and overview. In: Hutter, M. and Throsby, M., eds., *Beyond price value in culture, economics, and the arts*. Cambridge: Cambridge University Press. pp. 1-19.
- Hutter, M. and Throsby, M., eds., 2008. *Beyond price value in culture, economics, and the arts*. Cambridge: Cambridge University Press.
- ICOM, 2006. *Code of ethics for museums*. <<http://archives.icom.museum/ethics.html#intro>>, [accessed 24.05.13].

ICOM, 2012. *Museum definition*. <<http://icom.museum/the-vision/museum-definition/>>, [accessed 24.05.13].

IFLA, 2012. *Code of ethics for librarians and other information workers*. <<http://www.ifla.org/news/ifla-code-of-ethics-for-librarians-and-other-information-workers-full-version>>, [accessed 24.05.13].

International Council on Archives, 2011. *Draft manual on appraisal*. [online]. Available at: <<http://www.ica.org/11274/toolkits-guides-manuals-and-guidelines/draft-manual-on-appraisal.htm>>, [accessed 24.05.13].

International Organization for Standardization, 2003. ISO 14721:2003. *Space data and information transfer systems - open archival information system - reference model*. Geneva: International Organization for Standardization.

International Organization for Standardization, 2005. ISO 19005-1:2005. *Document management - electronic document file format for long-term preservation - part 1: use of PDF 1.4 (PDF/A-1)*. Geneva: International Organization for Standardization.

InterPARES2, 2008a. Appendix 16: Overview of the records continuum concept. In: Duranti, L. and Preston, R., eds., 2008. *International research on permanent authentic records in electronic systems (InterPARES) 2: experiential, interactive and dynamic records*. [online], Padova, Italy: Associazione Nazionale Archivistica Italiana. Available at: <<http://www.interpares.org/ip2/book.cfm>>, [accessed 24.05.13].

InterPARES2, 2008b. Project glossary. In: Duranti, L. and Preston, R., eds., 2008. *International research on permanent authentic records in electronic systems (InterPARES) 2: experiential, interactive and dynamic records*. [online], Padova, Italy: Associazione Nazionale Archivistica Italiana. Available at: <<http://www.interpares.org/ip2/book.cfm>>, [accessed 24.05.13].

Jenkinson, H., 1922. *A manual of archive administration*. Oxford: Clarendon Press.

Jenkinson, H., 1937. *A manual of archive administration*. revised ed. Oxford: Clarendon Press.

JISC, n.d.a. *NESLI2*. <http://www.jisc-collections.ac.uk/nesli2/>, [accessed 01.05.13].

JISC, n.d.b. *Digital-Preservation discussion list*. <https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=DIGITAL-PRESERVATION>>, [accessed 01.05.13].

JISC, 2009. *Establishing a digital preservation policy*.

<http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/establishing-a-digital-preservation-policy>

[accessed 01.05.13].

Johnson, P., and MacEwan, B., eds., 1993. *Collection management and development: Issues in an electronic age*. Chicago: American Library Association.

Johnson, P., 2009. *Fundamentals of collection development and management*. Chicago: American Library Association.

Jones, M. and Beagrie, N., 2001. *Preservation management of digital materials: a handbook*. London: British Library.

Jones, M. and Semple, N., 2006. Mind the gap: digital preservation needs in the UK. *Ariadne* [online], 48. Available at: <http://www.ariadne.ac.uk/issue48/semple-jones/>, [accessed 01.05.13].

Jones, S., 2009. *A report on the range of policies required for and related to digital curation*.

[online], Edinburgh: Digital Curation Centre. Available at:

[http://www.dcc.ac.uk/webfm\\_send/129](http://www.dcc.ac.uk/webfm_send/129), [accessed 01.05.13].

Jones, S., 2010. *DCC preservation policy template*. [online], Edinburgh: Digital Curation Centre.

Available at : <[http://www.dcc.ac.uk/webfm\\_send/236](http://www.dcc.ac.uk/webfm_send/236), [accessed 01.05.13].

Kalay, Y.E., Kvan, T. and Affleck, J., eds., 2008. *New heritage*. London: Routledge.

Katre, D., 2011. Digital preservation: converging and diverging factors of libraries, archives and museums - an Indian perspective. *IFLA Journal*, 37(3), pp.195-203.

Kenney, A.R. and McGovern, N.Y., 2005. The three-legged stool: institutional response to digital preservation (ppt.). [online]. March, 2005. II Convocatoriadel Coloquio de marzo, Cuba.

Available at: <[www.library.cornell.edu/iris/dpo/docs/Cuba-ark-nym\\_final.ppt](http://www.library.cornell.edu/iris/dpo/docs/Cuba-ark-nym_final.ppt)>, [accessed 24.05.13].

Knell, S.J., ed., 2004. *Museums and the future of collecting*. 2nd ed. Ashgate, Aldershot.

Knight, G. and Pennock, M., 2009. Data without meaning: establishing the significant properties of digital research. *International Journal of Digital Curation* [online]. 4(1). Available at:

<http://www.ijdc.net/index.php/ijdc/article/viewFile/110/87>, [accessed 24.05.13].

Kuhn, T., 1970. *The structure of scientific revolutions*. Chicago: University of Chicago Press.

- Kuipers, T. and van der Hoeven, J., 2009. *PARSE.Insight: Survey report*. [online]. Available at: <[www.parse-insight.eu/downloads/PARSE-Insight\\_D3-4\\_SurveyReport\\_final\\_hq.pdf](http://www.parse-insight.eu/downloads/PARSE-Insight_D3-4_SurveyReport_final_hq.pdf)>, [accessed 24.05.13].
- Kvale, S., 1996. *Interviews: an introduction to qualitative research interviewing*. Thousand Oaks; London: Sage.
- Ladkin, N., 2004. Collections management. In: Boylan, P., ed. *Running a museum: a practical handbook*. Paris: ICOM. pp.17-30.
- Lavoie, B., 2003. *The incentives to preserve digital materials: roles, scenarios, and economic decision-making*. [online], OCLC Research. Available at: <<https://www.oclc.org/content/dam/research/activities/digipres/incentives-dp.pdf>>, [accessed 24.05.13].
- Lavoie, B. and Dempsey, L., 2004. Thirteen ways of looking at...digital preservation. *D-Lib Magazine* [online]. 10(7). Available at: <<http://www.dlib.org/dlib/july04/lavoie/07lavoie.html>>, [accessed 01.05.13].
- Layder, D., 1993. *New strategies in social research: an introduction and guide*. Cambridge: Polity Press.
- Legal Deposit Libraries (Non-Print Works) Regulations 2013*. SI 2013/777. London: HMSO. Available at: <http://www.legislation.gov.uk/uksi/2013/777/contents/made>>, [accessed 01.05.13].
- Libraries and Archives Canada, 2007. *Digital collection development policy*. <http://www.collectionscanada.gc.ca/collection/003-200-e.html>, [accessed 01.05.13].
- Lincoln, Y.S. and Guba, E.G., 1985. *Naturalistic inquiry*. London: Sage Publications.
- Linstone, H.A. and Turoff, M., eds., 1975. *The Delphi method: techniques and applications*. [online]. Available at: <<http://is.njit.edu/pubs/delphibook/>>, [accessed 01.05.13].
- Little, J. and Naidoo, R., eds., 2005. *The politics of heritage: the legacies of 'race'*. London: Routledge.
- Lloyd, A., 2007. Guarding against collective amnesia? making significance problematic: an exploration of issues. *Library Trends* [online], 56(1), pp.53-65.<  
[http://muse.jhu.edu/journals/library\\_trends/v056/56.1lloyd.pdf](http://muse.jhu.edu/journals/library_trends/v056/56.1lloyd.pdf)>, [accessed 01.05.13].

LOCKSS, n.d. *What is LOCKSS?* <<http://www.lockss.org/about/what-is-lockss/>> [accessed 24.05.13].

London Metropolitan Archives, 2010. *Interim digital preservation policy*. [online]. Available at: <<http://217.154.230.218/NR/rdonlyres/6466F6FA-2F04-4E3E-8D8D-9158FD303425/0/DigitalPreservationPolicyJun2010.pdf>>, [accessed 01.05.13].

Lunghi, M., Grindley, N., Stoklasova, B., Trehub, A. And Egger, C., 2012. Economic alignment. In: McGovern, N. and Skinner, K., eds. *Aligning national approaches to digital preservation*. Atlanta: Educopia Institute Publications. pp. 195-234.

Malley, I., 1990. *Censorship and libraries*. London: Library Association.

Marchionni, P., 2009. Why are users so useful? User engagement and the experience of the JISC digitisation programme. *Ariadne* [online], 61. <http://www.ariadne.ac.uk/issue61/marchionni/>, [accessed 01.05.13].

Maron, N.L., Yun, J. and Pickle, S., 2013. *Sustaining our digital future: institutional strategies for digital content*. [online], Strategic Content Alliance. Available at: <<http://sca.jiscinvolve.org/wp/portfolio-items/sustaining-our-digital-future-institutional-strategies-for-digital-content/>>, [accessed 24.05.13].

Marstine, J., 2005. *New Museum theory and practice: an introduction*. Oxford: Blackwell.

Matthews, G., Poulter, A. and Blagg, E., 1997. *Preservation of digital materials policy and strategy issues for the UK: Report of a meeting held at the British Library Research and Innovation Centre, London, 13th December 1996*. British Library Research and Innovation Report 41. [online], British Library Research and Development Department. Available at: <<http://www.ukoln.ac.uk/services/papers/bl/blri041/digpres.html#contents>>, [accessed 01.05.13].

McGovern, N., 2007. *Version 2.0 Digital preservation policy framework: outline*. [online], ICPSR. Available at: <<http://www.icpsr.umich.edu/files/ICPSR/curation/preservation/policies/dp-policy-outline.pdf>>, [accessed 24.05.13].

McGovern, N. and Skinner, K., eds., 2012. *Aligning national approaches to digital preservation*. [online], Atlanta: Educopia Institute Publications. Available at: <[www.educoia.org/sites/educoia.org/files/ANADP\\_Educoia\\_2012.pdf](http://www.educoia.org/sites/educoia.org/files/ANADP_Educoia_2012.pdf)>, [accessed 24.05.13].

McInnes, S. And Phillips, V., 2009. *Survey of digital preservation in Wales summary of findings*. [online], Archives and Records Council Wales Digital Preservation Group. Available at: <[http://www.llgc.org.uk/fileadmin/documents/pdf/DP\\_Archives-Wales\\_Report-final.pdf](http://www.llgc.org.uk/fileadmin/documents/pdf/DP_Archives-Wales_Report-final.pdf)>, [accessed 01.05.13].

McKenzie Owen, J., 2005. Preserving the digital heritage: roles and responsibilities for heritage repositories. In: de Lusenet, Y. and Wintermans, V., eds. *Preserving the digital heritage principles and policies*. Amsterdam: Netherlands National Commission for UNESCO, pp. 45-49.

McLeod, R., Wheatley, P. and Ayris, P., 2006. *Lifecycle information for e-literature: full report from the LIFE project*. [online], London: LIFE Project. Available at: <<http://discovery.ucl.ac.uk/1854/>>, [accessed 24.05.13].

Meddings, C., 2011. Digital preservation: the library perspective. *The Serials Librarian*, 60(1), pp.55-60.

Mercer, H., 2004. *Appraisal policy background paper – the 'Grigg system' and beyond*. [online], London: The National Archives. Available at: [www.nationalarchives.gov.uk/documents/information-management/background\\_appraisal.pdf](http://www.nationalarchives.gov.uk/documents/information-management/background_appraisal.pdf), [accessed 24.05.13].

Meyer, E.T., Eccles, K., Thelwall, M. and Madsen, C., 2009. *Final Report, 20 July 2009 usage and impact study of jisc - funded phase 1 digitisation projects and the toolkit for the impact of digitised scholarly resources (TIDSR)*. [online], Oxford: Oxford Internet Institute. Available at: <[www.jisc.ac.uk/media/documents/programmes/digitisation/tidsr\\_finalreport.pdf](http://www.jisc.ac.uk/media/documents/programmes/digitisation/tidsr_finalreport.pdf)>, [accessed 24.05.13].

Mikolaskova, I., ed., 2010. *KEEP (Keeping Emulation Environments Portable) newsletter March 2010*. [online]. Available at: <[https://www.jiscmail.ac.uk/cgi-bin/webadmin?A3=ind1004&L=DIGITAL-PRESERVATION&E=base64&P=37656&B=-----cfbf54797fd167db4343a7ca309969be&T=application%2Fpdf;%20name=%22KEEP\\_Newsletter\\_March10.pdf%22&N=KEEP\\_Newsletter\\_March10.pdf&attachment=q&XSS=3](https://www.jiscmail.ac.uk/cgi-bin/webadmin?A3=ind1004&L=DIGITAL-PRESERVATION&E=base64&P=37656&B=-----cfbf54797fd167db4343a7ca309969be&T=application%2Fpdf;%20name=%22KEEP_Newsletter_March10.pdf%22&N=KEEP_Newsletter_March10.pdf&attachment=q&XSS=3)>, [accessed 24.05.13].

Miles, M.B. and Huberman, A.M., 1994. *Qualitative data analysis: an expanded sourcebook*. London: Sage Publications.

Moore, R., Jeffery, S. and Richards, J., 2010. *A review and analysis of the sustainability of digital curation and access to heritage data*. [online], CARARE. Available at:

<<http://www.carare.eu/eng/Resources/Review-and-Analysis-of-Sustainability-of-Digital-Curation-and-Access-to-Digital-Heritage>>, [accessed 01.05.13].

Morrow, T., Beagrie, N., Jones, M. and Chruszcz, J., 2008. *A comparative study of e-journal archiving solutions*. [online] London: JISC. Available at: <[www.jisc.ac.uk/media/documents/committees/jcs/26/jcs\\_08\\_20%20jisc%20collections%20-%20annex%20b%20e-journal%20archiving%20solutions%20final%20report.pdf](http://www.jisc.ac.uk/media/documents/committees/jcs/26/jcs_08_20%20jisc%20collections%20-%20annex%20b%20e-journal%20archiving%20solutions%20final%20report.pdf)>, [accessed 24.05.13].

Muir, A., 2004. Digital preservation: awareness, responsibility and rights issues. *Journal of Information Science*, 30(1),pp.73-92.

Muir, A., 2006. *Preservation, access and intellectual property rights challenges for libraries in the digital environment*. [online], London: IPPR. Available at: <<http://www.ippr.org.uk/publicationsandreports/publication.asp?id=464>>, [accessed 01.05.13].

Muir, A., Buttler, D. and Mossink, W., 2012. Legal Alignment. In: McGovern, N. and Skinner, K., eds. *Aligning national approaches to digital preservation*. Atlanta: Educopia Institute Publications. pp. 43-74.

Museums Association, 2004. *Acquisition guidance on the ethics and practicalities of acquisition*. [online], London: Museums Association. Available at: <<http://www.museumsassociation.org/download?id=11114>>, [accessed 01.05.13].

Museums Association, 2008. *Code of ethics*. [online], London: Museums Association. Available at: <[www.museumsassociation.org/download?id=15717](http://www.museumsassociation.org/download?id=15717)>, [accessed 24.05.13].

National Library of Australia, 2013. *Digital preservation policy 4th edition*. <http://www.nla.gov.au/policy-and-planning/digital-preservation-policy>>, [accessed 01.05.13].

National Library of Wales, 2008. *Digital preservation policy and strategy*. [online], National Library of Wales. Available at: <[http://www.llgc.org.uk/fileadmin/documents/pdf/2008\\_digipres.pdf](http://www.llgc.org.uk/fileadmin/documents/pdf/2008_digipres.pdf)>, [accessed 01.05.13].

National Museum of Australia, 2012. *Digital preservation and digitisation policy*. [online], National Museum of Australia. Available at: <[www.nma.gov.au/\\_\\_data/assets/pdf\\_file/0013/1453/POL-C-028\\_Digital\\_preservation\\_and\\_digitisation-2.2\\_public.pdf](http://www.nma.gov.au/__data/assets/pdf_file/0013/1453/POL-C-028_Digital_preservation_and_digitisation-2.2_public.pdf)>, [accessed 24.05.13].

- Neumayer, R. and Rauber, A., 2007. *Why appraisal is not 'utterly' useless and why it's not the way to go either*. [online], DPE. Available at: [http://www.digitalpreservationeurope.eu/publications/position/appraisal\\_final.pdf](http://www.digitalpreservationeurope.eu/publications/position/appraisal_final.pdf), [accessed 24.05.13].
- NINCH, 2003. *The NINCH guide to good practice in the digital representation and management of cultural heritage materials*. [online], Washington DC: National Initiative for a Networked Cultural Heritage. Available at: [www.nyu.edu/its/humanities/ninchguide/](http://www.nyu.edu/its/humanities/ninchguide/), [accessed 24.05.13].
- Northumberland Collections Service, 2007. *Digital preservation policy*. [online], Northumberland County Council Records Management Service. Available at: <[www.northumberland.gov.uk/idoc.ashx?docid=4f3d7b20-abe8-4e8f-932b-624851ca86d4&version=-1](http://www.northumberland.gov.uk/idoc.ashx?docid=4f3d7b20-abe8-4e8f-932b-624851ca86d4&version=-1)>, [accessed 24.05.13].
- Oliver, G., Chawner, B. and Liu, H., 2011. Implementing digital archives: issues of trust. *Archival Science*, 11(3/4), pp.311-327.
- Ooghe, B. and Moreels, D., 2009. Analysing selection for digitisation: current practices and common incentives. *D-Lib Magazine*. [online]. 15(9). Available at: <http://www.dlib.org/dlib/september09/ooghe/09ooghe.html>, [accessed 24.05.13].
- Owen, J.M., 2007. Preserving the digital heritage: roles and responsibilities for heritage repositories. In: de Lusenet, Y., and Wintermans, V., eds. *Preserving the digital heritage: Selected papers of the international conference*. Netherlands: National Commission for UNESCO. pp. 45-49.
- Padfield, T., 2010. *Copyright for archivists and records managers*. 4th ed. London: Facet Publishing.
- PARADIGM, 2008. *Workbook on Digital Private Papers*. <<http://www.paradigm.ac.uk/workbook/index.html>>, [accessed 24.05.13]
- Patel, M., Coles, S., Giaretta, D., Rankin, S. and McIlwrath, B., 2009. The role of OAIS representation information in the digital curation of crystallography data. In: Kellenberger, P., ed. *Fifth IEEE International conference on e-Science*. 9-11 December 2009. Los Alamitos, California: IEEE Computer Society. pp.132-139.
- Peach, C. and Foster, J., 2013. *Knowing the need optimising preservation for library and archive collections*. [online], London: British Library. Available at: [www.bl.uk/blpac/pdf/ktn.pdf](http://www.bl.uk/blpac/pdf/ktn.pdf), [accessed 24.05.13].

- Pennock, M. and Clark, C., 2011. Saving the sounds of the UK in the UK soundmap. *Ariadne*. [online]. 66. <http://www.ariadne.ac.uk/issue66/pennock-clark/>, [accessed 24.05.13].
- Pennock, M., 2013. *Web-archiving*. [online], York: DPC. Available at: <[dx.doi.org/10.7207/twr13-01](http://dx.doi.org/10.7207/twr13-01)>, [accessed 24.05.13].
- Phillips, M., 2009. PANDORA, Australia's web archive, and the digital archiving system that supports it. *National Library Of Australia Staff Papers*. Available at: <http://www.nla.gov.au/openpublish/index.php/nlasp/article/view/1249/1534>, [accessed 24.05.13].
- PORTICO, 2011. *Preservation of digitized books and other digital content held by cultural heritage organizations*. [online], PORTICO. <<http://www.portico.org/digital-preservation/wp-content/uploads/2010/01/NEH-IMLS-D-book-model.pdf>>, [accessed 24.05.13].
- PREMIS Editorial Committee, 2011. *PREMIS data dictionary for preservation metadata version 2.1*. [online]. Available at: [www.loc.gov/standards/premis/v2/premis-2-1.pdf](http://www.loc.gov/standards/premis/v2/premis-2-1.pdf)>, [accessed 24.05.13].
- Pymm, B., 2006. Building collections for all time: the issue of significance. *Australian Academic & Research Libraries* [online]. 37(1). pp. 61-73. Available at: <http://www.alia.org.au/publishing/aarl/37.1/collection.buildings.pdf>, [accessed 24.05.13].
- Reed, B., 2005. Challenges to born digital material. In: Gorman, G.E. and Shep, S.J., eds. *Preservation management for libraries, archives and museums*. London: Facet. pp.117-132.
- Reed, B., 2009. *Archival appraisal and acquisition*. In: *Encyclopedia of Library and Information Sciences*. 3rd ed. New York: Taylor and Francis. pp.120-129.
- Renaissance East Midlands, 2010. *Reviewing significance: a framework for assessing museum collections' significance, management and use*. [online], Renaissance East Midlands. Available at: <<http://www.collectionslink.org.uk/programmes/reviewing-significance/671-collections-reviewing-significance-framework>>, [accessed 24.05.13].
- Riestra, R., Beltran, X., Georgiou, P., Tsakonas, G., Kaur, K., Reilly, S., and Schmitt, K., 2013. *Business preparedness report*. [online], APARSEN. Available at: <<http://www.alliancepermanentaccess.org/wp-content/plugins/download-monitor/download.php?id=D36.1+Business+preparedness+report>>, [accessed 30.04.13].

RLG-NARA Digital Repository Certification Task Force. 2007. *Trustworthy repositories audit & certification: criteria and checklist*. [online], OCLC and CRL. Available at: [www.crl.edu/PDF/trac.pdf](http://www.crl.edu/PDF/trac.pdf), [accessed 24.05.13].

Robson, C., 2011. *Real world research*. Chichester: Wiley.

Rosenthal, D.S.H., Robertson, T., Lipkis, T., Reich, V. Morabito, S., 2005. Requirements for digital preservation systems a bottom-up approach. *D-Lib Magazine*. [online]. 11(11). Available at: < <http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html>>, [accessed 24.05.13].

Rosenthal, D.S.H., 2010. Format obsolescence: assessing the threat and the defenses. *Library Hi Tech*, 28(2), pp.195 - 210.

Rosenthal, D.S.H., Rosenthal, D.C., Miller, E.M., Adams, I.F., Storer, M.W. and Zadok, E., 2012. *The economics of long-term digital storage*. [online]. Available at: <[www.lockss.org/locksswp/wp-content/uploads/2012/09/unesco2012.pdf](http://www.lockss.org/locksswp/wp-content/uploads/2012/09/unesco2012.pdf)>, [accessed 24.05.13].

Ross, S., 2007. *Digital preservation, archival science and methodological foundations for digital libraries*. [online]. Keynote Address at the 11th European Conference on Digital Libraries (ECDL), Budapest (17 September 2007). <[http://www.ecdl2007.org/Keynote\\_ECDL2007\\_SROSS.pdf](http://www.ecdl2007.org/Keynote_ECDL2007_SROSS.pdf)>, [accessed 24.05.13].

Ross, S., 2012. digital preservation, archival science and methodological foundations for digital libraries. *New Review of Information Networking*, 17(1), pp.43-68.

Rothenberg, J., 1999. *Avoiding technological quicksand: finding a viable technical foundation for digital preservation: a report to the Council on Library and Information Resources*. Washington, DC: Council on Library and Information Resources.

Runardotter, M., Mortberg, C. and Mirijamdotter, A., 2011. The changing nature of archives: whose responsibility? *Electronic Journal of e-Government*. [online]. 9(1). pp. 68-78. Available at: [www.ejeg.com/issue/download.html?idArticle=224](http://www.ejeg.com/issue/download.html?idArticle=224)> [accessed 24.05.13].

Rusbridge, C., 2006. Excuse me ... some digital preservation fallacies? *Ariadne*. [online]. 46. Available at: <http://www.ariadne.ac.uk/issue46/rusbridge/>, [accessed 24.05.13].

Russell, K., 1999. *Why can't we preserve everything?* <<http://www.webarchive.org.uk/wayback/archive/20050410120000/http://www.leeds.ac.uk/cedars/colman/ABS01.html>>, [accessed 24.05.13].

Russell, R. and Winkworth, K., 2009. *Significance 2.0: a guide to assessing the significance of collections*. [online], Collections Council of Australia. Available at: <http://www.environment.gov.au/heritage/publications/significance2-0/pubs/significance20.pdf> [accessed 24.05.13].

Ruusalepp, R., Lee, C.A., van der Werf, B. and Woolard, M., 2012. Standards alignment. *In: McGovern, N. and Skinner, K., eds. Aligning national approaches to digital preservation*. Atlanta: Educopia Institute Publications. pp. 115-166.

Salza, S., Guercio, M., Grossi, M., Proll, S., Stroumboulis, C., Tzitzikas, Y., Doerr, M., Flouris, G. 2012. *Report on authenticity and plan for interoperable authenticity evaluation system*. [online]. APARSEN. Available at: [aparsen.digitalpreservation.eu/pub/Main/ApanDeliverables/APARSEN-DEL-D24\\_1-01-2\\_3.pdf](http://aparsen.digitalpreservation.eu/pub/Main/ApanDeliverables/APARSEN-DEL-D24_1-01-2_3.pdf), [accessed 24.05.13].

Samuel, R., 1994. *Theatres of memory vol.1 Past and present in contemporary culture*. London: Verso.

Sayer, A., 2000. *Realism and social science*. London: Sage.

Schellenberg, T.R., 1956. *Modern archives: principles and techniques*. Chicago: University of Chicago Press.

Seadle, M., 2004. Selection for digital preservation. *Library Hi Tech*. [online]. 22(2). pp. 119-121. Available at: <http://www.emeraldinsight.com/journals.htm?issn=0737-8831&volume=22&issue=2> [accessed 24.05.13].

Seadle, M., Rauber, A., Rusbridge, A., Schrimpf, S., and Schultz, M., 2012. Technical alignment. *In: McGovern, N. and Skinner, K., eds. Aligning national approaches to digital preservation*. Atlanta: Educopia Institute Publications. pp. 167-194.

Seadle, M., 2012. Archiving in the digital world: the scholarly literature. *Library Hi Tech*. [online]. 30 (2). pp.367 - 375. Available at: <http://www.emeraldinsight.com/journals.htm?articleid=17028094>>, [accessed 24.05.13].

Shen, L., Dorris Cassidy, E., Elmore, E., Griffin, G., Manolovitz, T., Martinez, M., and Turney, L.M., 2011. Head first into the patron-driven acquisition pool: a comparison of librarian selections versus patron purchases. *Journal of Electronic Resources Librarianship*, 23(3), pp. 203-218.

- Shepherd, E. And Yeo, G., 2003. *Managing records a handbook of principles and practice*. London: Facet Publishing.
- Silberman, N., 2008. Chasing the unicorn? The quest for “essence” in digital heritage. In: Kalay, Y.E., Kvan, T. and Affleck, J., eds. *New heritage*. London: Routledge, pp. 81-91.
- Silverman, D., ed., 2004. *Qualitative research: theory, method and practice*. London: Sage.
- Silverman, D., 2006. *Interpreting qualitative data: methods for analyzing talk, text and interaction*. 3rd ed. London: Sage.
- Simpson, D., 2005. *Digital preservation in the regions*. London: Museums, Libraries and Archives Council.
- Sinclair, P., Duckworth, J., Jardine, L., Keen, A., Sharpe, R., Billenness, C., Farquhar, A. and Humphreys, J., 2011. Are you ready? Assessing whether organisations are prepared for digital preservation. *International Journal of Digital Curation*. [online]. 6(1). pp. 268-281. Available at: <http://ijdc.net/index.php/ijdc/article/view/178/247>, [accessed 24.05.13].
- Smith, J.K. and Hodkinson, P., 2005. Relativism criteria and politics. In: Denzin, N.K. and Lincoln, Y.S., eds. *The Sage Handbook of Qualitative Research*. Thousand Oaks; London: Sage Publications, pp.915-932.
- Smith, L., 2006. *Uses of heritage*. Abingdon: Routledge.
- Smithies, R., 2011. *A review of research and literature on museums and libraries*. London: Arts Council England.
- State and University Library Denmark, 2012. *Digital preservation policy for the state and university library version 2.0*. [online]. Available at: <en.statsbiblioteket.dk/about-the-library/ddpolicy>, [accessed 24.05.13].
- State Library of Victoria, n.d. *Digital preservation policy*. < <http://www.slv.vic.gov.au/about-us/plans-policies/digital-preservation-policy>>, [accessed 24.05.13].
- Stawowczyk Long, A., 2009. *Long-term preservation of web archives – experimenting with emulation and migration methodologies*. [online], IIPC. Available at: <[www.netpreserve.org/sites/default/files/resources/Methodologies.pdf](http://www.netpreserve.org/sites/default/files/resources/Methodologies.pdf)>, [accessed 24.05.13].

- Strodl, S., Becker, C., Neumayer, R. and Rauber, A., 2007. *How to choose a digital preservation strategy: evaluating a preservation planning procedure*. [online]. Available at: <[www.ifs.tuwien.ac.at/~strodl/paper/FP060-strodl.pdf](http://www.ifs.tuwien.ac.at/~strodl/paper/FP060-strodl.pdf)>, [accessed 24.05.13].
- Strodl, S., Petrov, P. and Rauber, A., 2011. *Research on digital preservation within projects co-funded by the European Union in the ICT programme*. [online]. Available at: [www.ifs.tuwien.ac.at/%7Estrodl/paper/Report%20-%20Research%20on%20Digital%20Preservation.pdf](http://www.ifs.tuwien.ac.at/%7Estrodl/paper/Report%20-%20Research%20on%20Digital%20Preservation.pdf)>, [accessed 24.05.13].
- Sturges, P., 2003. Doing the right thing: professional ethics for information workers in Britain. *New Library World*, 104(1186), pp.94-102.
- Tanner, S., 2010. *Inspiring research, inspiring scholarship. The value and benefits of digitised resources for learning, teaching, research and enjoyment*. [online], London: JISC. Available at: <http://www.jisc.ac.uk/media/documents/programmes/digitisation/12pagefinaldocumentbenefitssynthesis.pdf>, [accessed 24.05.13].
- Tanner, S. and Deegan, M., 2011. *Inspiring research, inspiring scholarship. The value and benefits of digitised resources for learning, teaching, research and enjoyment*. [online], London: JISC. Available at: [www.kdcs.kcl.ac.uk/fileadmin/documents/Inspiring\\_Research\\_Inspiring\\_Scholarship\\_2011\\_SimonTanner.pdf](http://www.kdcs.kcl.ac.uk/fileadmin/documents/Inspiring_Research_Inspiring_Scholarship_2011_SimonTanner.pdf)>, [accessed 24.05.13].
- Tanner, S., 2012. *Measuring the impact of digital resources: the balanced value impact model*. [online]. Available at: <[www.kdcs.kcl.ac.uk/fileadmin/documents/pubs/BalancedValueImpactModel\\_SimonTanner\\_October2012.pdf](http://www.kdcs.kcl.ac.uk/fileadmin/documents/pubs/BalancedValueImpactModel_SimonTanner_October2012.pdf)>, [accessed 24.05.13].
- Taylor-Roe, J., 2009. 'To everything there is a season': reflections on the sustainability of the 'big deal' in the current economic climate. *Serials*, 22(2), 113-121.
- The National Archives, n.d. *Who we are*. <<http://www.nationalarchives.gov.uk/about%5Cwho-we-are.htm>>, [accessed 24.05.13].
- The National Archives, 2011a. *Archive principles and practice: an introduction to archives for non-archivists*. [online], London: The National Archives. Available at: <[www.nationalarchives.gov.uk/documents/information-management/archive-principles-and-practice-an-introduction-to-archives-for-non-archivists.pdf](http://www.nationalarchives.gov.uk/documents/information-management/archive-principles-and-practice-an-introduction-to-archives-for-non-archivists.pdf)>, [accessed 24.05.13].

- The National Archives, 2011b. *Digital preservation policies: guidance for archives*. [online], London: The National Archives. Available at: <http://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf> , [accessed 24.05.13].
- The National Archives, 2012a. *Appraisal policy*. [online], London: The National Archives. Available at: <[www.nationalarchives.gov.uk/documents/information-management/appraisal\\_policy.pdf](http://www.nationalarchives.gov.uk/documents/information-management/appraisal_policy.pdf)>, [accessed 24.05.13].
- The National Archives, 2012b. *Disposal scheduling*. [online], London: The National Archives. <[www.nationalarchives.gov.uk/documents/information-management/sched\\_disposal.pdf](http://www.nationalarchives.gov.uk/documents/information-management/sched_disposal.pdf)>, [accessed 24.05.13].
- The National Archives, 2012c. *Framework of standards*. [online], London: The National Archives. Available at: < <http://www.nationalarchives.gov.uk/archives-sector/standards.htm>>, [accessed 24.05.13].
- The National Archives, 2013. *Best practice guide to appraising and selecting records for The National Archives*. [online], London: TNA. Available at: [www.nationalarchives.gov.uk/documents/information-management/best-practice-guide-appraising-and-selecting.pdf](http://www.nationalarchives.gov.uk/documents/information-management/best-practice-guide-appraising-and-selecting.pdf), [accessed 24.05.13].
- The Scottish Government, n.d. *Ministers and law officers*. <<http://www.scottish.parliament.uk/msps/12445.aspx>>, [accessed 24.05.13].
- The Scottish Government, 2013. *Arts and culture*. <<http://www.scotland.gov.uk/Topics/ArtsCultureSport/arts>>, [accessed 24.05.13].
- Thibodeau, K., 2002. *Overview of technological approaches to digital preservation and challenges in coming years*. [online] <<http://www.clir.org/PUBS/reports/pub107/thibodeau.html>>, [accessed 24.05.13].
- Thomas, D.R., 2006. A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation* [online], 27(2). <<http://aje.sagepub.com/content/27/2/237.abstract#cited-by>>, [accessed 24.05.13].
- Thompson, M., 1979. *Rubbish theory: the creation and destruction of value*. New York: Oxford University Press.

Throsby, D. 1997. Seven questions in the economics of cultural heritage. *In*: Hutter, M. and Rizzo, I., eds. *Economic perspectives on cultural heritage*. Basingstoke: Macmillan. pp.13-30.

Throsby, D., 2001. *Economics and culture*. Cambridge: Cambridge University Press.

Todd, M., 2006. *Requirements for electronic records management systems functional requirements for the sustainability of electronic records*. [online]. London: The National Archives. Available at: <[www.nationalarchives.gov.uk/documents/functional\\_requirements.pdf](http://www.nationalarchives.gov.uk/documents/functional_requirements.pdf)>, [accessed 24.05.13].

Trant, J., 2009. Emerging convergence? Thoughts on museums, archives, libraries, and professional training. *Museum Management and Curatorship*, 24(4), pp. 369 - 387.

UK Data Archive, 2010. *UK Data archive preservation policy*. [online], UK Data Archive. Available at: <<http://www.data-archive.ac.uk/media/54776/ukda062-dps-preservationpolicy.pdf>>, [accessed 24.05.13].

UK Research Reserve, n.d. *Frequently asked questions*. <<http://www.ukrr.ac.uk/faq/default.aspx>>, [accessed 24.05.13].

ULCC and UKOLN, 2008. *The preservation of web resources handbook*. [online]. London: JISC. Available at: <http://www.jisc.ac.uk/media/documents/programmes/preservation/powrhandbookv1.pdf>>, [accessed 24.05.13].

UNESCO, 2003. *Charter on the preservation of digital heritage*. <[http://portal.unesco.org/en/ev.php-URL\\_ID=17721&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html)>, [accessed 24.05.13].

University of Edinburgh Library, 2010. *Preservation policy*. <<http://www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/data-repository/service-policies/preservation-policy>>, [accessed 24.05.13].

Upward, F., 1996. Structuring the records continuum - part one: postcustodial principles and properties. *Archives and Manuscripts*, 24(2), pp. 268-285.

Upward, F. 1997. Structuring the records continuum, part two: structuration theory and recordkeeping. *Archives and Manuscripts*, 25(1), pp. 10-35.

- Usherwood, B., Wilson, K. and Bryson, J., 2005. Perceptions of archives, libraries, and museums in modern Britain. *Library and Information Research*. [online]. 29(93). Available at: <<http://www.lirjournal.org.uk/lir/ojs/index.php/lir/article/view/202>>, [accessed 24.05.13].
- Verheul, I., 2006. *Networking for digital preservation: current practice in 15 national libraries*. IFLA Publications 119. [online], München: K.G. Saur. Available at: <[www.ifla.org/files/assets/hq/publications/ifla-publications-series-119.pdf](http://www.ifla.org/files/assets/hq/publications/ifla-publications-series-119.pdf)>, [accessed 24.05.13].
- Vermaaten, S., Lavoie, B. and Caplan, P., 2012. Identifying threats to successful digital preservation: the spot model for risk assessment. *D-Lib Magazine*. [online]. 18(9/10). Available at: <<http://www.dlib.org/dlib/september12/vermaaten/09vermaaten.html>>, [accessed 24.05.13].
- Waller, M. and Sharpe, R., 2006. *Mind the gap assessing digital preservation needs in the UK*. [online], London: Digital Preservation Coalition. Available at: [www.dpconline.org/docs/reports/uknamindthegap.pdf](http://www.dpconline.org/docs/reports/uknamindthegap.pdf), [accessed 24.05.13].
- Waters, D. and Garrett, J., 1996. *Preserving digital information report of the task force on archiving of digital information*. [online], Washington, DC: Council on Library and Information Resources. Available at: [www.clir.org/pubs/reports/pub63/reports/pub63watersgarrett.pdf](http://www.clir.org/pubs/reports/pub63/reports/pub63watersgarrett.pdf), [accessed 24.05.13].
- Welsh Government n.d.a. *John Griffiths AM*. <<http://wales.gov.uk/about/cabinet/cabinetm/johngriffiths;jsessionid=C534E8C4A07106096C112095AA8E0665?lang=en>>, [accessed 24.05.13].
- Welsh Government n.d.b. *CyMAL: Museums Archives and Libraries Wales*. <<http://wales.gov.uk/topics/cultureandsport/museumsarchiveslibraries/cymal;jsessionid=C534E8C4A07106096C112095AA8E0665?lang=en>>, [accessed 24.05.13].
- Wheatley, P. and Hole, B., 2009. *LIFE3: Predicting long term digital preservation costs*. [online]. In: iPRES 2009: the Sixth International Conference on Preservation of Digital Objects, San Francisco, USA, 5-6 October 2009, pp. 206-210. California Digital Library, UC Office of the President. Available at: <http://escholarship.org/uc/item/23b3225n#page-3>>, [accessed 24.05.13].
- Whyte, A. and Wilson, A., 2010. *How to appraise & select research data for curation*. [online], Digital Curation Centre and Australian National Data Service. Available at: [www.dcc.ac.uk/webfm\\_send/828](http://www.dcc.ac.uk/webfm_send/828), [accessed 24.05.13].

- Wilson, A., Wright, R., Polfreman, M., Anderson, S., Tanner, S. and Beer, E., 2006. *Digital moving images and sound archiving study*. [online], Arts and Humanities Data Service. Available at: <[www.jisc.ac.uk/media/documents/programmes/preservation/movingpicturesandsoundarchiving/finalversion.pdf](http://www.jisc.ac.uk/media/documents/programmes/preservation/movingpicturesandsoundarchiving/finalversion.pdf)>, [accessed 24.05.13].
- Wilson, A., 2007. InSPECT: *significant properties report*. [online]. Arts and Humanities Data Service. Available at: [www.significantproperties.org.uk/wp22\\_significant\\_properties.pdf](http://www.significantproperties.org.uk/wp22_significant_properties.pdf)>, [accessed 24.05.13].
- Winsor, P., 2009. PAS197: 2009. *Code of practice for cultural collections management*. [online], London: British Standards Institution. Available at: <<http://shop.bsigroup.com/en/ProductDetail/?pid=000000000030175180>>, [accessed 24.05.13].
- Xie, S.L., 2007. *Foundation for developing digital preservation policy: The InterPARES policy framework*. [online]. Available at: <[http://www.armaedfoundation.org/pdfs/iPRES2007\\_Beijing\\_Paper\\_SherryXie\\_Canada\\_20071011\\_.pdf](http://www.armaedfoundation.org/pdfs/iPRES2007_Beijing_Paper_SherryXie_Canada_20071011_.pdf)>, [accessed 24.05.13].
- Yale University Library, 2007. *Policy for the digital preservation*. [online], Yale University. Available at: <<https://collaborate.library.yale.edu/preservation/public%20Documents/Policies%20and%20Procedures/dig%20pres%20review%202011.pdf>>, [accessed 24.05.13].
- Young, L., 2004. Collecting: reclaiming the art, systematising the technique. In: Knell, S.J., ed. *Museums and the Future of Collecting*. 2nd ed. Ashgate, Aldershot. pp.185-195.
- Youngs, K., 2001. *Managing the digitisation of library, archive and museum materials*. [online], London: National Preservation Office. Available at: <http://www.bl.uk/blpac/pdf/digitisation.pdf> [accessed 24.05.13].
- Zorich, D., Waibel, G. and Erway, R., 2008. *Beyond the silos of the LAMs*. [online]. Dublin, Ohio: OCLC Research. Available at: <[www.oclc.org/research/publications/library/2008/2008-05.pdf](http://www.oclc.org/research/publications/library/2008/2008-05.pdf)>, [accessed 24.05.13].

## APPENDICES

---

### Appendix 1

Compiled from Beagrie et al 2008, p.16-17 and p.24-25.

1. Principle Statement: mission, Need for policy, benefits to organisation
2. Contextual Links: how it integrates with other policies
3. Preservation Objectives: high level purpose, aim of preservation actions
4. Identification of Content: outline what the policy's overall scope is in terms of content and its relationship to collection development aims.
5. Procedural Accountability: Identify high level responsibilities for the policy and provide recognition of the most important obligations faced in preserving key institutional resources.
6. Guidance and Implementation - how the policy will be implemented, including the following:
7. Financial and Staff Responsibility: who is responsible for digital preservation within the organisation, high level and staff; how digital preservation fits into financial plans
8. Intellectual Property: how this will be managed; legal context, access, deposit agreements etc
9. Distributed Services: whether anything is outsourced
10. Standards Compliance: standards the organisation is committed to
11. Review and Certification: how often the policy is to be reviewed
12. Auditing and Risk Assessment: how and what risks assessed e.g. file formats, legal framework, audit trails, exit strategy
13. Stakeholders: who is involved in the policy and its implementation
14. Preservation Strategies: technical procedures for preservation
15. Glossary; Version control: history and bibliographic details of the version.

## Appendix 2

Table 18 About the policy guides

Guide	Context of guide (who created it?)	Function of guide (what is it supposed to be for?)	Audience of guide (who is it aimed at?)
Erpanet Digital Preservation Policy tool 2003	European-funded project team	Part of a range of tools developed by project	Institutions, especially cultural heritage
CHIN 2006	CHIN is part of the Department of Canadian Heritage	As part of a series of reference resources aimed at enabling museum professionals	Museums primarily
IPSCR model 2007	Nancy McGovern – director of digital preservation at IPSCR. Organisation is a large international consortium of academic institutions and research organisations.	Proposes a digital preservation policy framework for any organisation to use to develop their own policy framework.	Primarily member organisations, i.e. research intensive organisations, but available freely on their website
Beagrie et al 2008	Funded by JISC, focuses on the UK Higher and Further Education sectors. Authors are external consultants, from Charles Beagrie Limited. Based on an analysis of a range of existing policies and guides.	To help higher and further education institutions formulate policy – gives both a recommended set of clauses and shows how a digital preservation policy relates to other internal policies	Higher and further education institutions
JISC 2009	JISC Digital Media (formerly TASI) are based in the University of Bristol. They exist to provide support and advice to higher and further education institutions	It is a brief overview document giving practical advice as to how to formulate a policy	Aimed at those with responsibility for managing digital media collections within higher and further education
DCC Preservation Policy Template 2010	The DCC are an organisation primarily serving the Higher Education sector. They promote good practice and knowledge sharing, and provide information and training	To assist in the definition of a digital preservation policy; gives brief description of clauses with examples from the four policies and guides they used to formulate this guide, which were from the archival/repository sector	Institutional repositories
TNA guidance 2011	The National Archives are the official national archive for the UK Government and provide leadership and guidance top other archival institutions	To help organisations improve governance of digital material through the development of a policy	Primarily publically funded archives though also other archives

### Appendix 3

Table 19 Comparison of policy guides

	ERPANET 2003	CHIN 2006	IPSCR 2007	Beagrie et al 2008	JISC 2009	DCC 2010	TNA 2011
Principle statement	Y	Y	Y	Y	Y	Y	Y
Contextual links	Y	Y	Y	Y	Y	N	Y
Preservation objectives	Y	Y	Y	Y	Y	Y	Y
Identification of content	N	Y	Y	Y	Y	Y	Y
Procedural accountability	Y	N	Y	Y	Y	N	Y
Financial and staff responsibility	Y	N	Y	Y	Y	N	Y
Intellectual property	Y	Y	Y	Y	N	N	Y
Distributed services	Y	N	Y	Y	N	N	Y
Standards compliance	Y	Y	Y	Y	N	Y	Y
Review and certification	Y	Y	Y	Y	N	N	Y
Auditing and risk assessment	Y	N	Y	Y	N	N	Y
Stakeholders	N	N	Y	Y	N	N	Y
Preservation strategies	Y brief	Y	Y	Y	Y	Y	Y
Glossary	N	Y	Y	Y	Y	N	Y
Process for accessing records	N	N	Y	N	N	N	Y
Withdrawal of collections	N	N	N	N	N	Y	N
Closure of service	N	N	N	Y	N	Y	N
Disaster/Emergency plan	N	N	N	Y	N	N	N
Value of material	Y	N	N	N	N	N	N
Ownership/responsibility	N	Y	N	N	N	N	N
Selection/acquisition	N	Y	Y	N	N	N	N

## Appendix 4

Table 20 Practitioner interview schedule

Interview Questions	Further notes and follow up questions
What does your job involve? What types of material do you deal with?	How does the overall mission of the parent organisation influence the work of your institution?
Do you have a preservation policy for either analogue or digital material?	How was it devised? By whom? Was any particular guidance followed? Who is responsible for it?
Do you have policies or procedures for deleting or de-accessioning material?	How is deletion or de-accessioning approached in your institution?
How is material selected?	Who in your institution makes these decisions or has input into making them?
What factors influence selection decision making in your institution?	Factors could include such things as costs, expertise, opportunity etc.
What criteria are used for selection?	How were these devised? Who by? Which are most important?
Is digital and traditional material treated differently?	Why?
What aspects of the material are you trying to preserve when you take preservation actions?	This refers to 'significant properties', so for example some institutions may be more interested in preserving the data rather than the media.
How do you determine the value (not necessarily financial) of objects or collections?	How can you deal with potential changes in value?
Are there any legal or ethical influences on your selection and preservation work?	How do they affect your activities?
Who are your users?	What influence do they have on selection (directly or indirectly)?
Much digital material is outside of institutions so who do you think is, or should be, responsible for preserving digital material in general?	What role do you think institutions have in selecting and/or preserving material? Are there roles for individual practitioners and individuals outside institutions?
What do you think might be important influences on selection and preservation in the future?	
What do you think are the necessary skills to select and preserve material?	Do you have to appropriate skills and resources in your institution? What training do you think is needed?

**Appendix 5**

**Table 21 Stakeholders**

all stakeholders	archive	libraries	museums	examples
creators	x	x	x	developers, artists, authors
external users	x	x	x	The public, family history researchers
other organisations, including parent organisations	x	x	x	local council
local informants or enthusiasts	x		x	Community groups
other cultural heritage institutions	x		x	
national institutions	x	x		BL, TNA
volunteers	x		x	
commercial companies and the mass market	x	x		Suppliers, publishers, software vendors
external funders		x	x	DCMS, HLF, charities, arts councils, JISC, research councils
colleagues	x	x	x	collection development/acquisitions staff
practitioners	x	x	x	
internal users	x	x	x	
senior managers or management team	x	x		
selection committees	x		x	
IT dept/systems section	x	x		

**Appendix 6**

**Table 22 Comparison of policies and policy guides**

	National Museum of Australia	UK Data Archive	ADS	EROS Centre	West Yorks Archives	Northumberland Collections Service	Gtr Manchester Archives	Parliamentary Archives	London Metropolitan Archives	Hampshire Archives
preservation objectives	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
principle statement	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
contextual links	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
identification of content	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
preservation strategies	Y	Y	Y	Y	Y	Y	Y	Y		Y
standards compliance	Y	Y	Y	Y	Y	Y		Y	Y	Y
glossary; version control	Y	Y	Y	Y	Y	Y			Y	Y
Financial and staff responsibility	Y	Y	Y			Y				
review and certification	Y	Y	Y		Y	Y	Y	Y		N*
auditing and risk assessment		Y	Y			Y		Y		
procedural accountability	Y	Y	Y		Y	Y			Y	

process for accessing records		Y	Y		Y	Y	Y			Y
guidance and implementation	Y		Y	Y				Y	Y	
distributed services	Y			Y	Y			Y	Y	
challenges of digital material	Y							Y		Y
sources for policy guidance	Y			Y				Y		
stakeholders	Y		Y	Y					Y	
disaster/Emergency plan						Y		Y		Y
requirements for acceptance of deposits		Y	Y							
intellectual property			Y							
sustainability or closure of service										
withdrawal of collections		Y								

	Libraries Archives Canada	NLNZ/ ANZ	National Library of Australia	National Library of Wales	RCAHMS	Royal Denmark Library	Wellcome Library	Yale	State & University Library Denmark	State of Victoria Library	Dspace	University of Edinburgh repository
preservation objectives	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
principle statement	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N*	
contextual links	Y	Y	Y	Y		Y	Y	Y	Y	Y		
identification of content	Y	Y	Y	Y		N*	Y	Y	Y	Y	Y	
preservation strategies		Y	Y			N*	Y	Y	Y	Y	Y	Y
standards compliance	Y		Y	Y		Y	Y	Y	Y			
glossary; version control		Y				Y	Y	Y	Y			Y
financial and staff responsibility	Y	Y		Y			Y	Y	Y	Y		
review and certification		Y		N*		Y	Y		Y			
auditing and risk assessment	Y	Y		Y	Y			Y	Y			
procedural accountability	Y		Y						Y			
process for accessing records	N*	N*				Y			Y		Y	

Guidance and implementation	Y			Y				Y				
distributed services	Y		Y									
challenges of digital material	Y		Y			Y		Y				
sources for policy guidance					Y	Y		Y				Y
stakeholders	Y	Y	N*					N*	Y*	N*		
disaster/Emergency plan					Y		Y					
requirements for acceptance of deposits							Y				Y	
intellectual property							Y				Y	
sustainability or closure of service					Y						Y	Y
withdrawal of collections											Y	Y

