Stakeholders in the selection of digital material for preservation: relationships, responsibilities, and influence

This item was submitted to Loughborough University’s Institutional Repository by the/an author.


Additional Information:

- This is an Accepted Manuscript of an article published by Taylor & Francis in Collection Management on 07 May 2015, available online: http://dx.doi.org/10.1080/01462679.2015.1011816

Metadata Record: https://dspace.lboro.ac.uk/2134/18299

Version: Accepted for publication

Publisher: © Routledge (Taylor & Francis)

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Stakeholders in the Selection of Digital Material for Preservation: Relationships, Responsibilities and Influence

CLARE RAVENWOOD, ADRIENNE MUIR, and GRAHAM MATTHEWS
Loughborough University, Loughborough, United Kingdom

Selecting digital material for preservation in libraries, archives and museums is a necessary task but has not been widely examined, although the nature of digital material challenges traditional methods of selecting. This paper examines the social context of selection in institutions, in which the responsibilities of stakeholders and relationships between them can affect the material chosen for preservation by practitioners. A range of stakeholders is identified; relationships between practitioners, IT staff and sources of material are found to be crucial. The influence of senior managers is important in providing a mandate and encouraging shared working and networks of expertise.

KEYWORDS: digital preservation, digital materials, archivists, information technology,
INTRODUCTION

Many libraries, archives and museums have become increasingly engaged with digital preservation as their collections have included more digital material. Material that was traditionally collected in hard copy is now produced and used in electronic formats and much of this needs to be preserved for the future. 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 selecting digital material for preservation. This paper focuses on selection and in particular the stakeholders in the selection process, their responsibilities which impact selection and the ways in which relationships between stakeholders affect selection practice.

‘Digital material’ is either digital versions of analogue material that has been ‘digitised’ to make a copy, or ‘born digital’ material that does not have an analogue equivalent (Digital Preservation Coalition (DPC) 2008, 24). This includes not only digital versions of books or journals, but also other types of material such as images, sound files, websites and multimedia objects. Preserving digital material is important to libraries, archives and museums as it enables them fulfill an ongoing heritage role through collecting digital material for the benefit of current and future users. Heritage involves passing on knowledge to future generations; in order to do this it must be recorded and preserved (Gorman 2007, 286) which is often the responsibility of libraries, archives and museums (Feather 1996, 58; Usherwood, Wilson and Bryman 2005). Digital preservation supports broader activities, for example through enabling continuing delivery of services or the ongoing use of digital tools to increase user interaction and access in museums (Hudson 2012, 44).
As Beagrie, Lavoie and Woollard (2010, 56) highlight in a discussion of the benefits of digital preservation, it is not the process or the set of preserved information that is important, but the value of the activities performed with the preserved material.

Digital material has particular properties and requirements different to those of non-digital material, which means a different approach to preservation is needed. Digital preservation can be defined as ‘the series of managed activities necessary to ensure continued access to digital materials for as long as necessary’ (DPC 2008, 24). Similarly the American Library Association (ALA) (2009) provides the following definition: ‘Digital preservation combines policies, strategies and actions that ensure access to digital content over time’. Both these definitions emphasise that digital preservation is an on-going activity that requires considerable investment of time and resources; it is a continuing process, not a one off action, and includes policies and strategies in combination with actions. This is different to preserving non-digital material, which can be preserved ‘passively’ by providing the appropriate storage environment and controlling environmental variables such as temperature, humidity and pests (Adcock 1998) without the need for an active program of preservation to ensure it is continually accessible and useable.

Selection for digital preservation is seen as expensive (Blue Ribbon Task Force (BRTF) 2010, 46; Lunghi et al 2012, 218), yet necessary (DPC 2008, 103), and it is unclear whether traditional processes or strategies are effective in selecting digital material. Different types of memory institutions have different approaches to selection, based on contrasting origins, traditions and roles (Trant 2009, 370). Briefly, 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, 125). 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, 46) but the role of the individual curator has traditionally been to use their own judgment when selecting what to collect (Young 2004, 189). Archival appraisal is based strongly within a traditional document and organisational context and has a body of underpinning theory on which practice is based. Appraisal involves making a judgment on the continuing value of material, so material can be discarded that has no further value (Eastwood 2004, 202; TNA 2013, 15) and is most often associated with making value decisions prior to acceptance of responsibility (Craig 2004, 44). Traditionally the role of the archivist is to be an advisor and a protector of records (Duranti 1994, 343). 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, 124; Cook 2011, 177). Libraries on the other hand focus more on collecting for current use and providing access to material for their users. Whilst the mission of libraries, similarly to museums and archives, may have a preservation element (Maron, Yun and Pickle 2013, 36), not all library material is collected with an assumption of permanence; the material which is kept permanently tends to be found in ‘special collections’ in libraries (Feather 2004, 13). Looking after digital material in the long term in libraries is more analogous to the activities and role of archives and archivists (Feather 2004, 12; Ross 2012, 50).

While selection is little explored, the views of practitioners (meaning those who have a professional interest in or responsibility for collections in institutions, such as librarians, archivists and curators) on the question of selection have been
under-explored in particular. Future cultural heritage is at least in part dependent on the selection decisions made now within institutions by practitioners; the selection for preservation process creates heritage, as criteria determine what is to be valued and defined as heritage (Cameron 2008, 177). Practitioners’ expectations and experiences of change due to increasing amounts of digital material have not yet been considered, though they are the people performing selection on behalf of their institution.

This paper reports on research that was carried out as part of a project on selection for digital preservation. This part of the project had two aims: to identify stakeholders that may have an interest in or influence on selection for preservation in libraries, archives and museums and to investigate their roles and responsibilities. Whilst it is important to gain an overview of practitioners they 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. Through in-depth interviews with digital preservation experts and practitioners from a range of libraries, archives and museums it became clear that other stakeholders have a part to play in selection; it is an activity performed in a social context wherein others can influence selection choices. However, the roles and responsibilities of other stakeholders and their impact on selection decisions have up to now been overlooked. This paper identifies those stakeholders that have an influence on selection, and it examines the role of stakeholders by exploring their responsibilities which may affect selection practice.
LITERATURE

Digital Preservation

There are a number of differences between digital and non-digital material that make the preservation of digital material more complex. Digital material depends on changing technology, where both hardware and software can become obsolete; changes to formats, coding, operating systems and other aspects can render information unusable. Much digital preservation literature focuses on the risk to data from software format obsolescence (Rothenberg 1998; Waller and Sharpe 2006, 8; DPC 2008, 36; del Pozo, Stawowczyk Long and Pearson 2010, 292 for example); however this focus has been questioned. Rusbridge (2006) and Rosenthal (2010) ask whether formats are changing as quickly and irrevocably as is assumed. There are other, perhaps more pressing, threats to digital material than format obsolescence; Rosenthal et al (2005) and Barateiro et al (2010, 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; economic and organisational failure; and legal changes and requirements. Digital media is fragile; magnetic devices such as hard drives can fail mechanically and are vulnerable to a range of environmental factors, as are optical storage devices such as CDs and DVDs, which also have a short expected lifespan (DPC 2008, 154). There is an enormous amount of digital material created every year and institutional digital collections can be large; the APARSEN project (Riestra et al 2012, 14) surveyed 101 European research libraries, including government, university and national libraries, about their preparedness for digital preservation and found that 49% held digital collections greater than 10Tb in size. The form of digital heritage material may vary widely; a European survey of cultural heritage institutions
(Stroeker and Vogels 2012, 15) found that born-digital material in collections included:

‘photographs, video's /DVD/CD, other audiovisual objects or audio-recordings, TV and radio programmes, film, archives and archival records, E-books and E-journals, web pages and websites, computer software and games… datasets, interviews, oral histories and PDF (or other formats)’.

Digital material is global, it may originate from anywhere and geographic boundaries may not apply (Owen 2007, 46). Digital material is dynamic and interactive; it can be constituted through social interaction, such as social media or online games, and is linked in multiple ways through networks. Text, sound and images are no longer separate information types stored on different media, but they can be combined into complex multimedia products. These products can incorporate different works with different origins, so the legal issues of ownership of rights can be complex (Muir et al 2012; Charlesworth 2012; Anderson 2013, 18). These factors, along with the need for an active program of preservation to ensure it is continually accessible and useable (DPC 2008, 25), mean that digital material is complex to maintain and keep useable over time.

Selection for Digital Preservation

Selection is viewed as necessary throughout the digital preservation literature as part of the process of preservation (DPC 2008, 103) for a number of reasons. It is assumed that the volume of digital information that is available to be preserved is too large for it to be feasible to preserve everything (for example Feeney 1999, 11; Deegan and Tanner 2006, 15; Harvey 2007, 9; Ooghe and Moreels 2009) although this view has been challenged, particularly on the basis of cost, where economies of
scale (Bearman 2007, 35) and potentially cheap storage and the use of search tools (Bailey 2008, 100) would mean that more could be kept. Related to this is the question of whether it would be wise to save everything; a further reason suggested for selection is quality control (Harvey 2005, 55), where only the ‘best’ material should be preserved. Another argument for selection pervasive throughout the literature is that there are inadequate resources to preserve all digital information, so selection is necessary (for example Russell 1999; Harvey 2007, 8; Berger 2009, 60), although this view has also been challenged (Rusbridge 2006). Such an assertion is based on an institutional model where preservation is only performed by libraries or archives for example, neglecting the contribution of others such as enthusiasts (Barwick, Dearnley and Muir 2011, 375). However, from a pragmatic perspective, the DPC Handbook (2008, 42) asserts that:

‘The enormous quantity of information being produced digitally, its variable quality, and the resource constraints on those taking responsibility to preserve long-term access, makes selectivity inevitable if the objective is to preserve ongoing access.’

Whilst the issue of resources is not limited to digital material, it becomes pressing when factors such as the cost of digital storage, expertise in information technology, creation of adequate metadata, and the need for constant interventions to keep material useable is considered. Sustainable business models are still unclear or untested (Evens and Hauttekeete 2011, 160; Ravenwood, Matthews and Muir 2013, 296) and 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, 11).
Digital material has particular requirements and properties meaning traditional forms of selection or appraisal may not be appropriate (Russell 1999; Eastwood 2004, 202; Bailey 2008, 72 for example). These include an increased volume of material, multiple copies or versions of the same item, and being stored on older media so the material needs processing first (Paradigm 2008). The NDSA report on the appraisal and selection of geospatial data also notes specific issues, including the use of commercial or proprietary formats, the complexity of databases and combinations of datasets, as well as the number and size of files (Morris 2013, 8-9). Therefore 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. Furthermore, as the 4C Roadmap (4C Project 2014, 6) explains, selection may need to become at least in part an automated process that is decided by algorithms due to the large volume of material, which is too great for traditional methods of selection. As is made plain by Lunghi et al (2012, 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.’

This may reflect the different traditions and approaches found in different libraries, archives and museums which are then continued for digital material. For example, few criteria have been formulated specifically for preservation of digital material; there may be an assumption is that the same criteria will be appropriate as that used
for traditional material or for digitisation. Although many of these may still be relevant, as Edwards, Matthews and Nankivell (2000, 21) note, selection for digitisation has different motivations and implications than selection for preservation. Harvey (2007, 10) 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. There is a need to include not only current high use but also potential future use as a criterion (Lunghi et al 2012, 218; Morris 2013, 12). 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. Many of these are similar to criteria used in selecting analogue material, but documentation and metadata assume greater importance for digital material; this is noted by Morris (2013, 12) as a common criterion. An overview of the literature and discussion of some of the issues in selection for digital preservation can be found in Ravenwood, Matthews and Muir (2013), including a comparison of criteria for digitisation and preservation. There is little else in the literature examining criteria specifically for preservation, or more recently for selection in general as selection is performed rather than analysed, but what there is highlights the importance of an institution having the necessary ability to preserve digital material, both technologically and economically (Seadle 2004; Whyte and Wilson 2010).

Stakeholders

Digital material, like traditional material, has multiple individual stakeholders: creators, publishers, rights holders, librarians, archivists, curators and users for
example, which means some form of cooperation is necessary for preservation. At different points in its lifecycle there may be a range of stakeholders that have responsibility for digital material; different stakeholders have influence on and interest in preservation at different stages (Lavoie and Dempsey 2004; DPC 2008, 65). 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) though progress has previously been slow in integrating this responsibility into institutional structures. The ‘Mind the Gap’ report (Waller and Sharpe 2006, 16), which surveyed individuals with an interest in digital preservation from many types of organisation in the UK, including libraries, archives, government bodies, research institutions and industry, found only 33% of organisations surveyed had a clear responsibility structure for digital preservation. A lack of clear responsibility was identified in the BRTF report (2009, 13) as a particular barrier to the successful implementation of digital preservation strategies in an organisation and this continues; the APARSEN Business Preparedness Report (Riestra et al 2012, 37) refers to ‘the current fragmentation of roles and responsibilities of the key stakeholders’. 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. Verheul (2006, 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, 31). A similar need for co-operation between groups was described by Runardotter et al
(2011, 76) who examined responsibility in public sector archives, where they found that ‘cooperation, coordination and communication’ between archivists, IT (information technology) 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 at the same time their influence in the organisation is limited.

The relationship of archivists and record managers with IT staff has been little explored, except on a small scale by Oliver, Chawner and Lui (2011), who surveyed IT professionals working in New Zealand government organisations for their opinions on digital archives and the role of record keepers. They concluded that firstly archivists and IT staff are not working together and secondly that IT staff view record keepers are 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 found to be very wide (2011, 321), echoing findings from Craig (2004, 75) who identified some of the potential difficulties encountered by archivists when faced with new forms of technology, including technical stakeholders that do not understand the archival point of view. Responsibility within an institution for preservation can be unclear if it is seen as ‘an IT problem’ (Jones and Semple 2006) neglecting the valuable input of information and curatorial professionals, especially in selection of material. The division between curatorial staff and technical staff does not seem to have changed since then; more recently Seadle et al (2012, 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’. 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), especially as there is as yet no clear ‘digital curator’ or ‘curation practitioners’ (4C 2014, 4) role in many institutions, as practitioners do not yet have the required technical IT skills.

Despite the need for librarians, archivists and curators to share responsibility with IT staff, it seems that a lack of common understanding can lead to a lack of ability to work together, which could jeopardise organisational efforts to preserve digital material. The need for interpersonal or collaborative skills training has been noted as important for digital preservation (Pryor and Donnelly 2009, 166; Gregory and Guss 2011, 186) but the provision of this within digital preservation curricula is lacking. Analysis of training provision was undertaken as part of the APARSEN project which found that training on ‘roles and responsibilities’, the most relevant aspect to collaborative working, was provided in only 21 of 134 initiatives (Cirinna, McMeekin and Kilbride 2012, 34). It is the development of these skills that would improve the ability of different stakeholder groups to work together and increase understanding.

The responsibility for selection for digital preservation can be unclear, as the term ‘selection’, along with similar terms such as appraisal or acquisition, has specific meanings depending on the context in which it are used. This relates to the traditional practices and theoretical perspectives of libraries, archives and museums. Separate educational regimes for librarians, archivists and curators have been suggested as a factor which hinders working together to manage digital material, as
these emphasise differences rather than similarities (Katre 2011; Trant 2009; Given and McTavish 2010, 22); this may make cross-institutional comparison and working more difficult (Ooghe and Moreels 2009). However, libraries, archives and museums face similar problems and issues regarding digital material (Marty 2014, 615) and therefore training in digital preservation which crosses traditional boundaries, such as from the DPC and ULCC (University of London Computer Centre), may help to overcome differences.

METHODOLOGY

The overall aim of the project was to investigate the theory and practice of selection for digital preservation in UK libraries, archives and museums. It was decided at an early stage that this research would focus on digital preservation in the UK, in order to make it manageable, but 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. This research consisted firstly of a review of the literature and then in depth open interviews with eight digital preservation experts were conducted to orientate the research. Six of them were from the UK, one from Canada and one from Australia.

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. They 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.
Next a series of twenty five semi-structured interviews were conducted 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. Eleven were from archives, five from museums and nine from libraries. They worked for a range of institutions, including national and local libraries, archives and museums of different sizes, data services, universities and two company archives. The latter are not typical stewardship organisations but it was felt useful to include a wide range of practitioners who have an interest in the preservation of digital material. Potential participants were found through a variety of means, including:

- MLA designated collections list
- Society of Archivists
- Culture24 listings of institutions
- JISC digital media website
- Other interviewees
- Respondents to a request for policy documents on JISC mailing lists

The initial sample was identified through a preliminary analysis of potential stakeholders to identify those who may have an interest in the topic, and later sampling focussed on the people with direct day-to-day involvement with the issues raised in the previous interviews. A qualitative approach was taken because when the context and individuals’ accounts are important, as in this research, then qualitative methods are more appropriate to gather and analyse the richness of data required. Interviews were an appropriate method of data collection as the objectives refer to opinions, perceptions and
assumptions as well as facts, which can easily be gathered through interviews with respondents.

FINDINGS AND DISCUSSION

The findings in this section are drawn in particular from the interviews with the experts (labelled E then a number) and the practitioners (P then a number). It became clear from the interviews that issues of selection for digital preservation span traditionally different communities of practitioners in libraries, archives and museums, despite previous arguments that differences in terminology and the education of librarians, archivists and curators are divisive (for example Given and McTavish 2010, 22; Ooghe and Moreels 2009).

When practitioners were questioned about selection, a range of stakeholders with an influence emerged from the interviews. These were:

- Practitioners
- Colleagues
- Managers
- Senior managers
- The institution
- Users
- Creators and donors (including authors, publishers and colleagues)
- IT staff
- External funders
- Other institutions
The practitioners discussed the activities and responsibilities of other stakeholders which impact on selection; issues of stakeholder responsibility were found to be broadly similar across all three traditionally separate types of institution. Figure 1 summarises the responsibilities of stakeholders that have a particular impact upon selection. Institutional 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.
Figure 1: stakeholder responsibilities
The Institution and Senior Managers

The practitioners viewed the institution as a whole as having very broad responsibilities that impact on selection. When discussing the institution and senior managers, the practitioners commonly conflated discussion of selection with digital preservation, reflecting their broader roles in managing both. Many of the practitioners were clear that it was the role of their institutions to preserve all types of material that 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 - manager at a large library)

The assumption of permanent responsibility once an item is in the collection continues with digital as with non-digital material. Institutional responsibilities included: caring for collections (regardless of media); working with stakeholders; having clear policies and procedures; providing support for staff; and working within legal requirements. The latter includes for example 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, Schedule1) which outline ways in which personal data should be managed responsibly by data controllers, including 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. These provide an overall structure for selection and the mission of the archive, which practitioners are expected to understand when making decisions.
The authority of the institution and practitioner in selecting material was reflected in the creation and use of preservation or selection policies as described by respondents. A policy provides direction and guidance to an internal audience and defines why an organisation is doing digital preservation, for both internal and external stakeholders, acting as the authority for those undertaking digital preservation (TNA 2011b, p.5). Respondents agreed with this, viewing policies as allowing practitioners to make decisions and giving room for them to use their judgment in what to select (P5; P13). None of the respondents reported that their institution had a policy specifically for selection of digital material however, although half of the practitioners said they had a written preservation policy for either analogue or digital. It seems that in some institutions materials to be selected are considered as part of a broader collecting policy which applies to both digital and on-digital material. P14 for example stated ‘we have our collections policy, our collecting policy, and of course the nature of a collecting policy is that it is medium-blind.’ The implication here is that there is little recognition in policy documents of the issues inherent in selecting specifically digital material.

The next layer toward the centre of figure 1 represents senior managers. Senior managers influence selection by setting boundaries to collection, achieved in a number of ways. They make clear the aim or mission of the institution. Senior managers formulate high level criteria; the experts differentiated between the responsibility for setting criteria, and the responsibility of practitioners for applying those criteria. Preservation or selection policies provide guidance for collecting and prioritisation and many practitioners considered 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 also limit selection by controlling and authorising 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.

The practitioners regarded senior management ‘buy in’, or support, for selection of digital material as particularly important; similarly it was considered necessary for supporting digital preservation. It was felt they should provide a mandate for preserving digital material, and hence support for selecting material to be preserved, whether as a new activity or as an extension of an existing mandate to preserve paper-based material. This provides a rationale for practitioners to use when dealing with other stakeholders, giving them the authority to select for preservation, and encourages an organisation wide engagement with digital preservation. These findings are consistent with conclusions from the BRTF report (2009, 23), which states a clear mandate is necessary even with incentives in place for digital preservation, and from the APARSEN report on Business Preparedness (Riestra et al 2012, 23) which found that the majority of respondent organisations thought that mandates were facilitators to digital preservation, including encouraging collaboration between different stakeholders. Evidence from this research supports this; examples were given by respondents who wanted to engage further with digital material but were limited in their ability to do so due to a lack of senior management support. For example P4 (a company archivist) controls the paper based archive; she stated that she had no senior management ‘champion’ and therefore had little influence over how the organisation selects and manages its electronic records,
which for historical reasons is controlled by the IT department. These findings are consistent with conclusions from Hudson (2012, 46) and the DPC Handbook (DPC 2008, 20). In contrast, P16 described how she was able to obtain a strong mandate and support for the development of an institutional repository in her organisation, which meant that other staff were obliged to engage with selection and preservation of material in the repository. Whilst the influence of senior managers on selection may be indirect, through high level policy or mandates for example, without their support and ‘championing’ for digital preservation it seems less likely to happen or be made more difficult, so the efforts of individuals in institutions to select material with a view to preservation will have less impact.

Operational Managers and Colleagues
The next layer towards the middle of figure 1 represents operational levels of management, some of whom took part in the practitioner interviews. Findings indicate a generic set of responsibilities which impact on both selection and digital preservation; these are not related to media, instead reflecting good management practice within an institution. For example making sure that staff have access to the appropriate training and there are adequately trained staff within the institution is seen as a management responsibility. Whilst this is generic it is perhaps more pressing with digital material as technical skills and knowledge may be needed to select effectively. The responsibility which was highlighted by practitioners 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 a need to keep up to date with digital preservation
issues by managers, and in particular to ‘share best practice’, perhaps through engagement with the digital preservation literature, appropriate JISC mailing lists or DPC events and news, which they would then be able to pass on to practitioners. This would benefit them both for selection and digital preservation.

Although the responsibilities identified above are similar to those with traditional selection, the evidence suggests that the social aspect of stakeholders as guides to selection is important for digital material. 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 can help to determine the value of material. According to respondents an important role of colleagues is to give informal support to selection decision making; ‘some items may be more ‘tricky’ to make a decision about and so colleagues give extra help’ (P7). 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. Practitioners were very clear how important and valuable formal and informal networks of expertise are to their selection practice. 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 … 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.’

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 (Angevaare et al 2012, 95; Portico 2011, 15) but here the emphasis is on the value of individuals being able to collaborate informally. Immediate colleagues are particularly important, and their influence on the practitioner performing selection is high by providing extra guidance and reassurance about difficult decisions. There seems to be great use of both online and face to face support networks by practitioners who engage with digital material, so these should be further developed and made more accessible. Although networks of digital preservation expertise exist, through the efforts of the DPC and DCC for example, the reach of such networks may not be far enough as some practitioners demonstrated a lack of awareness of sources of guidance. They also lacked the opportunity to consult colleagues or had 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 access such support.

IT Staff

IT staff in particular were discussed by respondents in relation to responsibilities for digital material. The interaction between curatorial staff and IT staff arises from the latter’s responsibility to manage the technical infrastructure and in some institutions to manage the material itself. In figure 1 IT staff are illustrated as part of the institution; 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. IT expertise may be acquired externally by the institution or be provided by internal staff, with whom the practitioners may have not previously had much contact. IT staff may be dedicated to the library or archive if it is a large institution, such as a national library or a data service, but are more likely to be responsible for a number of departments, where the memory institution is part of a larger organisation such as a university. The role of IT or systems staff is largely unacknowledged in the literature, as noted by Oliver, Chawner and Lui (2011, 314). Whilst IT staff may not have responsibility for selection, they may have responsibility for preservation and therefore have influence over what is kept. If they are not working with practitioners then the practitioners are not able to source material (suggested by P14), or be confident that what is selected will be safe (suggested by P4), or are not provided with the appropriate infrastructure and tools to both select and preserve digital material.

It is clearly an important relationship, specific to 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 as they felt they ‘owned’ any digital material. They did not recognise her role for collecting and preserving it, and she had a difficult relationship with them which blocked constructive conversation to resolve problems. Unclear roles and poor quality of relationships with IT staff may mean that the practitioner may have less control over the material, and workflows to process the material may not be understood. P19 (a manager at a large university library) 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 is a lack of a shared understanding of the function of an archival collection, illustrating well the conclusion from Oliver, Chawner and Lui (2011, 321) that the cultural differences between IT people and record managers are very wide. Many respondents were noticeably keen to differentiate themselves from others; P5 clearly stated ‘we are historians, we’re not technical people’. IT personnel are seen as having very different views of the material and of appropriate processes and procedures. It is here that the potential for the positive influence of senior managers can be seen, in providing not only a mandate for preservation but also a requirement for different areas of the institution to work together. Encouragement from senior managers for co-operative working and participation in events or training which include both IT staff and curatorial staff would help to improve understanding. There are few training courses that are aimed at both groups although, for instance, the DPTP (Digital Preservation Training Programme) from ULCC (University of London Computer Centre) is aimed at information management professionals, but the website also states that IT staff ‘should’ attend (http://www.dptp.org/). Much of the education for digital preservation or curation in the UK is offered by universities as postgraduate courses, as listed by the DigCurV (Digital Curator Vocational Education Europe) website (http://www.digcur-education.org/eng/Training-opportunities), which is not practical for most currently working with digital material, whether practitioner or IT staff. Attending vocational training courses is also not always practical either for a number of reasons, such as for lone practitioners (as described by P7 for example), because the institution
will not allow more than one person to attend at one time or due to travel costs (Arthur 2009, 4).

Despite possible difficulties with bringing practitioners and IT staff together in formal training situations, it may benefit institutions to develop collaborative understanding or relations between staff in other ways. Collaborations between different institutional departments may be informal or formal (Verheul 2006, 31). Despite some acknowledgement of the potential benefits of formal collaborations, the practitioners demonstrated a reluctance to engage in this with other stakeholders for digital material, especially IT staff, without clear boundaries and formal arrangements to provide the necessary structure. Some animosity was revealed in the practitioner interviews, which implies that working together needs careful management. Furthermore, practitioners felt that informal collaboration between individual stakeholders is equally important as formal collaborative activities. For instance, practitioners suggested that most training is done ‘on the job’, often through informal methods. Informal collaborations through networks which include both practitioners and IT staff may be an effective method for spreading awareness and knowledge, and staff should be encouraged to participate. This would help to bridge divides between different groups of stakeholders.

Practitioners
The inner circle nearest selection represents practitioners. Selection is seen as central to their view of themselves as professionals and this is the same regardless of media. It was at the personal level that practitioners were able to discuss selection in depth as it is often part of their day to day role. Examination of the expert and
practitioner interviews makes clear how in their view responsibilities will change very little specifically for digital material; they are often generic. In comparing responses from practitioners from different libraries, archives and museums 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.

The authority of the practitioner in selecting material continued to be assumed throughout. Selection in libraries in particular focuses on knowledge of the user as central to selection decision making (Clayton and Gorman 2001, 4; Harvey 2005, 58; Johnson 2009, 108) and many practitioners, not only librarians, claimed this is only possible through skills developed in their training. They implied 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, 180) and Smith (2006, 12) that heritage is circumscribed by ‘experts’. An example of this is where P12 questions the use of volunteers in archives. In her view they are not able to be dispassionate about the material and she implies that they might allow their personal preferences to influence their selection decisions, undermining the role of the archivist as a collector not creator of archives (Duranti 1994, 343). This seems short sighted and dismissive as volunteers may be professionals themselves. This conceptualisation of selection as solely a professional activity may become problematic as there could be tension between this and the need to automate selection (4C Project 2014, 6).
The experts agreed with the practitioners that selection is an activity performed by professional practitioners acting on behalf of current or future 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 scope for users to be collaborators in selection is circumscribed, and the traditional role of a practitioner as a ‘proxy’ when selecting material does not change with digital. 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.’

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 both now and in the future.

A key finding from the interviews that negative relationships with other stakeholders may engender a feeling of anxiety in practitioners from a perceived loss of control over selection. To them this means that there will be further committees, along with others’ priorities and views to take into account, which will slow and
complicate the process. A university archivist (P20) expressed worries that in the future for digital material selection will involve ‘so many other people than the archivist’ and will not be based on the archivist’s judgement. Conversely when considering broader issues of digital preservation, some respondents agreed that the responsibility could be shared with other stakeholders; this was seen 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. There is a continuum of contact between stakeholders (Zorich, Waibel and Erway 2008; Higgins 2012, 611) from an understanding of the professional landscape through co-operation, collaboration and ultimately convergence. Factors such as similar problems and issues in digital preservation (Marty 2014, 615; Higgins 2012); common guidance or codes of practice for collection management such as PAS197; and generic digital preservation training and education that is not aimed at a particular type of institution (Trant 2009, 377) can drive collaboration and ultimately convergence between different types of institution (Higgins 2012). Benefits accrue at each stage, such as opportunities to share resources and best practice (Higgins 2012, 611) but practitioners are not always aware of these, nor are they always ready to transcend traditional professional boundaries (Stewart 2012, 276). Cirinna, McMeekin and Kilbride (2012, 50) reported on focus groups of cultural heritage practitioners and managers undertaken as part of the DigCurV project on training needs; a recurring theme was the need to ‘Improve communication between technical and non-technical staff through a shared vocabulary’ (2012, 50). Therefore it is important for managers to create harmonious relationships, building trust and a shared understanding to reduce anxiety and encourage effective co-operation.
Users, Creators and Donors

Users, creators and donors are roles that other stakeholders both internal and external to the institution may have, which includes 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, even though both experts and practitioners clearly understood that it is impossible to guess the uses that future users may make of material selected now.

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 which asked the public to contribute their memories of World War Two (BBC 2012). 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 libraries, archives and museums is to act on their behalf. This is confirmed by this research; the evidence suggests that the question of ‘who has responsibility for selecting digital material?’ has the answer ‘the same people who do it for non-digital material’. This seems short sighted; it does not take advantage of the potential for increased engagement of users that digital material may afford nor take account of the networked nature of digital material that may allow different models of ownership or access.
Donors and creators have roles as sources of material, which influences selection in a number of ways. Firstly they determine the set of material that a selection is made from, either through providing material, by identifying material to be preserved, or by donating material. Practitioners described a range of sources, including publishers, artists, authors, parent organisations, other institutions and also the public, which echoes a suggested role of private individuals from the BRTF report (2010, 39). Positive relationships with sources can impact upon selection. Local informants and enthusiasts increase the availability of material to choose from as they alert the practitioner 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. Furthermore a responsibility of creators and donors from the view of the practitioners seems to be to make selection easier by not only bringing material to their attention but also using appropriate file formats. Sources limit selection through their choice of format, where use of an obscure format means it is less likely that the material will be selected; the institution may not be able to manage it or it may need a great deal of processing prior to preservation actions. Practitioners who accept digital donations described making preservation easier by influencing creators before deposit to use ‘preferred’ formats and to supply appropriate documentation or contextual information for their material (P8). The creator is encouraged to go through the selection process themselves before deposit 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.’ There is a need to build relationships with sources of material that do not just focus on
obtaining donations, but also giving influencing and providing guidance to creators or donors on the types of material that can be selected for preservation.

Sources can be problematic though, for example archivists described problems acquiring records from a parent organisation, where distant or difficult relationships limit collecting even where they have a responsibility for preserving the records of the organisation. Also the value of digital material can be affected through the ownership and management of rights by sources; value is closely linked to the ability to use of material (Ross 2007, 3). It is of less value and less likely to be selected if the institution is unable to use it now or to preserve it for later use due to rights issues. This is also linked to the quality or quantity of contextual information that sources supply, which may also affect whether the material is of value (stated by E4 and P10 for example), reflecting the suggested criterion of ‘documentation’ from the DPC Decision Tree (2006).

Not collecting contextual information could affect the ability to find or use the material and also the ability to determine authenticity. Authenticity is an archival concept meaning that the record is what it purports to be (Bearman and Trant 1998; Wilson 2007, 4); the DPC Handbook (2008, 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 material, whether paper or electronic, is important especially where it is used as evidence, but particularly so for digital where it may be changed easily. As Eastwood (2004, 205) makes clear, if there is doubt about the authenticity of records then they are less likely to be selected. 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 link between acquiring appropriate metadata and file formats from creators and the ability of the institution to provide a better preservation service is made explicit in the National Digital Stewardship Alliance (NDSA) ‘Levels of Digital Preservation’ guidelines (Phillips et al, 2013, 5). Good relationships with creators and donors allow practitioners to gather contextual data and practitioners suggested that this should be part of the acceptance procedure. Conversely 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. 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 where it is possible for them 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. The conceptualisation of archivist and record keeper’s roles in the more recent records continuum model (Upward 1996) is more flexible, recognising that record and archive keeping is a shared responsibility and that the archivist’s role is not only at the end of a linear process. This model was developed partly in response to the introduction of electronic records, and it seems to reflect practice as described here. Archival practitioners suggested that the way forward is to provide clear selection criteria and procedures in selection policies, ensuring both creators and practitioners understand the boundaries of selection and helping to apply these consistently.
CONCLUSIONS

The value of this study lies in its consideration of selection for digital preservation from the perspective of the practitioner, whose role and influence has been overlooked in the literature. They clearly see selection as their responsibility and as part of their role as a professional, but there is a disparity between their desire to retain selecting as a professional activity 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 and decision making more easily. Practitioners in libraries, archives and museums need to work more closely with other stakeholders and differences in approach need to be managed, especially between technical and curatorial staff where shared understanding is crucial. Training courses that include different stakeholder groups would be useful in not only teaching all the same standards and methods but also increasing understanding and improving the ability to work together, as would the encouragement of both formal and informal networks of expertise. The responsibility of management to support digital preservation activity, including selection, and enable staff to acquire appropriate skills or employ specialist staff is highlighted here.

This research supports previous assertions that selection is a social activity, partly through identifying a range of stakeholders that have an influence on selection decision making. They are more numerous and wide ranging than for selection of non-digital material, including technical IT staff in particular. 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. This is turn affects the value of the material and the ability of the institution to
successfully preserve it.

Relationships are a key factor in selection decision making and the quality of
relationships can either be useful or hinder selection. This concept was emphasised
throughout the interviews with experts and practitioners and the findings support the
argument from Lavoie and Dempsey (2004) that selection for digital preservation is a social
and cultural process. It should be viewed within a broader context than the individual; the
quality and form of relationships practitioners have with other stakeholders and the roles
that they play are important factors in selection; by encouraging relationships and networks
the ability and engagement of practitioners to select digital material effectively could be
increased.
REFERENCES


Harvey, Ross. 2007. *DCC Digital Curation Manual Instalment on Appraisal and Selection*. HATII, University of Glasgow; University of Edinburgh; UKOLN, University of Bath; Council for the Central Laboratory of the Research Councils.

http://www.dcc.ac.uk/webfm_send/121.


http://www.digitalpreservation.gov/ndsa/activities/levels.html


