

Subjectivity in Digitization
Joshua Chalifour
McGill University
GLIS 647 Independent Study
Supervisor: Dr. Eun Park
15 December 2014
DOI: 10.5281/zenodo.1312201

Table of Contents

Introduction.....	3
Literature Review.....	4
Methodology.....	12
Results.....	14
Discussion and Recommendations.....	23
Conclusion, Limitations, and Further Study.....	35
References.....	38
Appendices of Participant Commentary.....	41

Introduction

When information professionals such as archivists digitize their holdings, they enable access to content—they create digital objects that previously did not exist. People access these digital objects with a perception that does not necessarily reflect an understanding of the rationale for why the objects were created. The creation of these digital objects results from a selection process that is usually based on clear and concrete objective factors, however subjectivity with its more opaque factors, also influences selection. Currently, we do not account for the impact of subjectivity during the selection process but if we hope to provide people with a full understanding of what they're accessing, we ought to.

The purpose of this study was first to elicit information from professionals that engaged in digitization projects about the ways in which they made selection judgments. Second, it was to see whether there were common themes in their experiences, which could be given some sort of formal recognition. It investigates subjectivity in an effort to both complement existing guidance on how archivists themselves influence their digitization selections; and to make this subjectivity more transparent.

The traditional archival impression of neutrality and impartiality imbues many guides on digitization, which focus on objective factors to select materials for digitization. A number of these guides also state outright or hint at the fact that subjectivity influences the selection phase of digitization. Still, the influence of subjectivity on selection and how to address that influence in what the user receives, does not seem to have been fully explored. This suggests some professional uncertainty about how to gauge the subjective role of the archivist in what gets selected for digitization. The archivist or team that undertook the digitization process had reasons to create the object but these are neither always obvious nor transparent to the people accessing it.

In selecting records or a fonds for digital preservation, archivists must consider many factors that support the ongoing maintenance of, and access to the resulting digital object¹ for a potentially very long time so it's important that they think about the way future users intellectually receive what has been selected.

The following addresses the existing literature in terms of guidelines, subjectivity and the archivist's professional responsibilities, and the impact of subjectivity; then explains the study's methodology and the results. The results lend themselves to some recommendations for addressing the impact of subjectivity.

Literature Review

Guidelines

A number of guides or best practices documents exist (often from government or cultural institutions) such as the National Initiative for a Networked Cultural Heritage's (NINCH) guide², The Canadian Heritage Information Network's (CHIN) guide on planning and implementing digitization projects³, the Parliamentary Archives of the UK's guidelines⁴, and the digital preservation guidelines developed for the European Commission⁵. These detail step-by-step selection practices for digitization. Many guides (e.g. CHIN) focus primarily on objective selection factors. Some, for example the NINCH guide, recognize that certain factors (e.g. intellectual value) involve subjective judgments. In a section devoted to the process of selection, the guide presents the following ten questions to ask before for selecting material to digitize⁶.

1. Is the information content of the analog objects high?
2. What is the intellectual significance of the analog materials?
3. How do they compare in terms of importance, authority, uniqueness, and timeliness?
4. Is the object you are considering for digitization a good example of the period under examination?
5. Is it original?
6. Are there alternatives in the collection more capable of illustrating particular points?

7. Is it complete or are there any parts missing?
8. Is it up to date?
9. Is it accurate?
10. Would it withstand the test of time or is it only of ephemeral interest today? (Is it, for example, an important testimony of the past, of another culture or our own, of an artistic movement, an author, or a scientific discovery?)

The guide states that perceptions or evaluation can change so it recommends consulting with various stakeholders (peers, users, etc.) and seeking general consensus.

The scope of the selection process, in which we're concerned is described in the handbook from the Northeast Document Conservation Center (NEDCC) on managing “digital conversion projects from Maxine Sitts⁷. It moves from nominating materials to evaluating the nominated materials, and then prioritizing the remaining materials. This process includes factors such as assessing the ongoing maintenance of the fonds with respect to the institution’s mission and collection scope; legal requirements; value in terms of authenticity, effort, audience, etc.; and physical and technical concerns among other criteria.

Guidance such as the declaration of thirteen principles on digitization from the Canadian Council of Archives (CCA) helps archivists approach the process⁸. It states concrete verifications against existing institutional policies and provides few overarching statements to influence how archivists should conduct their activities toward digitization. While it primarily deals with objective factors, it neither explicitly excludes nor explicitly raises subjective ones—the guide says “Records should be chosen for digitization only after a careful selection process. They must meet project objectives...”⁹ thus if the project objectives are met with some dependency on the archivist's subjectivity, that could fall within the scope of these principles.

The US Federal Agencies Digitization Guidelines Initiative developed a planning and management outline that plots out a digitization workflow¹⁰ and offers guidance in assessment and prioritization focused on usage restrictions (legal) and approval processes. The guide

recommends that nominations should be based on institutional priorities and “research/public interest”¹¹ which, while not defined might presumably be based on frequent usage or some judgment as to the value of the content for the audience.

Preparing Collections for Digitization by Bülow, Ahmon, and Spencer details best-practices for digitization initiatives. The authors identify some primary selection factors and once these are satisfied, recommend other selection factors that “...can be distilled down to three core factors: content, demand and condition.”¹² They identify means of assessing these factors, which include getting knowledge from the institution's staff and inviting the public or user groups to nominate materials. Due to the potentially subjective locus these nominations come from, the authors recommend that a panel considers the proposals in order to get as many different perspectives as possible¹³.

In the NEDCC handbook on managing digital conversion projects, Diane Vogt-O'Connor likens selection to the archivists' appraisal process because it involves a decision on acquisition, prioritizing for conservation, and a purpose of exhibiting or publishing materials¹⁴. She discusses the appropriateness of the appraisal choices and recommends evaluating issues around content that would be ethically or culturally sensitive to stakeholders.¹⁵ One could certainly consider this a subjective judgment. Vogt O'Connor assumes the archivist will be burdened with judgments due to ethical and cultural sensitivities in his or her stance on what degree it is appropriate to actively participate in the public sphere.

Value, usually intellectual, according to guides like NINCH¹⁶ might be used during selection. A UNESCO-commissioned guide recommends that each project formulate a definition of intellectual value inline with its own goals¹⁷. The basis for this definition lies in the idea that potential users value the content enough that the effort and resources to digitize it were

worthwhile. Although archivists can identify value there is much difference in the way value can be determined. Thomas James Connors argues that the concept of valuation isn't well-defined¹⁸. Hazen, Horrel, and Merrill-Oldham's insights on identifying intellectual value call out a number of areas in which subjectivity influences selection. They argue that due to institutions' generally restricted resources, it is important to first figure out the intellectual value of the materials to be digitized¹⁹. They explain how the archivist judges the intellectual value of the fonds²⁰ based on information content and significance, whether the fonds is a “good” exemplar of its content, its ephemeral status within the context of other potential materials for preservation, and the likelihood that it merits ongoing accessibility to users.

The National Archives of the UK created a policy, which includes seven principles guiding the ongoing care of its collections²¹. The guidance does not aim to be a step-by-step procedure for digitization but offers perspective on value. It instructs archivists to consider the value of materials before selecting them for digitization. It breaks value into specific qualities²² that include the value of information (in any medium), material (that can represent age, provenance, or other information about the physicality), and culture (divided into things such as political, aesthetic, and ethical).

Lorna Hughes explains that archivists (among others) must be judicious in their work and organize their selections in accord with the strategy and policy of their institutions. In her book, *Digitizing Collections: Strategic Issues for the Information Manager*, she recommends developing selection policies before embarking on digitization. These policies will be influenced by a “...focus on the nature and intellectual content of the collections, their condition, and usage...”²³ Hughes calls out the necessity of determining whether the information content and intellectual value of the materials are worth digitizing.²⁴ She explains that archivists need to think

about the way their target user groups are likely to use the digital objects. This is not necessarily straightforward to figure out since the very act of providing a digitized object could increase or decrease usage of the fonds. While she recommends that selection criteria covering information content and intellectual value should be inline with the institutional mission and collections policy, she doesn't offer details on how determine intellectual value, nor what constitutes that value.

Subjectivity and the Archivist's Professional Responsibilities

The CCA defines digitization as “...the transformation of analog information (from whatever form and from whatever support) to digital code.”²⁵ The process of undertaking this transformation involves traditional archival responsibilities along with new ones, specific to the digital realm.

Joan Schwartz and Terry Cook observed that “Inside the profession, archivists have perceived themselves as neutral, objective, impartial.”²⁶ Jenkinson opposed personal judgments because he felt it would be detrimental to keeping an archive's reputation of impartiality²⁷. These perspectives contrast with “Archival Science and Postmodernism” in which Terry Cook raises informative points for an archivist's stance on selection (even though it was not written to address selection processes). Cook says “...archives are a sacred public trust of preserving society’s memories that must be widely shared. Archivists serve society, not the state...”²⁸ Cook writes that archivists put their own values into their work, which will necessitate a close look at the processes in building archives.²⁹ Accepting Cook's characterization suggests that archivists have a responsibility to analyze how their own judgments of society's memories stand in relation to serving society, which contrasts with the traditional perception of impartiality.

Hazen, Horrel, and Merrill-Oldham position archivists with the responsibility of arbiters

of scholarly value, arguing that if there is “marginal scholarly value”, not to bother with digitizing the material³⁰.

Mark Roosa discusses developments in attitudes toward preservation programs (including some digital techniques). He argues that preservation professionals have many new abilities they need to develop in order to address future problems,³¹ including the philosophical. Roosa discusses the US Library of Congress's copy of the Gutenberg Bible. Its digitized version opened students, researchers, and others to a new understanding of the significance that the object represents. It helped educate what sort of impact Gutenberg's invention had on our civilization's history. This is due in part to the accessibility of the digital object. The digital object creates new awareness and visibility into topics related to the object and not just the object itself. The archivist is yoked with a responsibility for considering this impact upon what he or she selects; Roosa says “Above all, preservation programs should be driving forces that responsibly guide and coordinate the art, craft, and science of preservation.”³²

In an article on selection for preservation in the digital age, Janet Gertz talks about the increase in work that digitization involves. She states “Using digitization to create access to materials of long-term value calls for genuine commitment to preserving the digital files, to development of the infrastructure necessary to preserve those files routinely over the long term.”³³ Thus, judgments during selection result in not just what users receive but also bond archivists to ongoing responsibilities.

The Impact of Subjectivity

Archivists interpret the content of their holdings and by contributing selections for the context of digitized presentations, shape some of the discourse around the subject matter. In a digitization project, selected materials can later be presented in contexts that influence how users

perceive the content. The result could profoundly affect the way researchers come to understand their subjects³⁴. It is not just that the selected fonds become available and accessible for research but due to the scarcity of time and resources, selecting certain fonds could preclude the selection of others (putting them at risk of not being available to future users³⁵) or present a challenge to understanding the content outside of the fonds' full context³⁶.

Examples of this are clear in Tom Nesmith's argument that the public doesn't realize how large a role archives play in major public affairs.³⁷ Archivists, knowing that people want to access digital information, are selecting content that helps to provide context for popular subject matter accessed digitally. Nesmith does not offer guidance on how archivists should make these judgments. In "Why We Collect: Curators, Collectors, and the Urge to Acquire", Douglas P. McElrath shows that there is a real act of influence on the public sphere, which comes from the archivist's choices. McElrath makes a case that what an archivist selects to digitize (and thus make accessible) from our history, increases peoples' awareness of that part of history.

Zhang and Gourley discuss how selection may be focused on materials with specific content because there is a need to satisfy a subject-oriented digital collection. They comment that this is commonly seen in collaborative digitization projects where each collection will be the source of part of the digitized materials³⁸. However, this requires that an archivist judge content relevancy and suggests that he or she anticipate specific realms of research or usage that will apply to the digitized object (both activities involve subjectivity). Thus the goal of creating something for a purpose that was not inherent to the original materials, influences the selection process.

Robert Cole and Chris Hackett in their essay on research in the digital environment look at user expectations for what is intellectually receivable and what is actually provided in a digital

context.³⁹ They refer to the great new accessibility that is possible by selecting many sources for digitization versus the difficulty of finding information from within those sources. Joshua Sternfeld points out the creative and personal role that archivists play in selecting records of historical events⁴⁰ where what is selected, itself shapes what users will be able to search and have context for. Due to end user research and search behaviour, as well as expectations on the availability of digitized materials, archivists need to account for how the resulting digital object they've decided to produce, will be perceived.

In Summary

The literature identifies numerous areas in which archivists' subjectivity influences the selection of materials for digitization. It also identifies that what is selected for digitization impacts the perception people have of the digital content upon access it. Existing literature offers great detail addressing objective factors. Yet, where subjectivity is involved it raises questions about how we approach the selection process. It does not resolve how we deal with the impact of subjectivity.

Methodology

This study focused on subjective aspects of the digitization process. Specifically, there were three main questions for inquiry.

- 1) What steps are taken in the digital preservation selection processes that archivists adopt in practice?
- 2) How do archivists perceive the influence their personal roles (including experience, attitudes, and skills) have in creating the intellectual object that results from a digital preservation selection process?
- 3) What are the recurring subjective factors in the identified processes that archivists make in their selections for digital preservation?

The research targeted the role and function in undertaking digitization projects, not

merely the archivist title. People can be employed under a different title yet do digitization work within an archival or similar sort of context. Thus, participants were targeted based on the type of work they had engaged in. Note that in discussing information from the people in this study, the terminology “participant” has been used. When talking about the profession, the term “archivist” has been used for simplicity (though it could be a librarian or other professional).

Ultimately, seven participants from six different institutions were chosen through three methods. First, through judgments on participants with probable experiences at local archival institutions including universities, museums, and community cultural organizations. Second, a request for participants was posted to a Web-based (LinkedIn) forum for archivists. Third, some snowball sampling techniques were used to find participants. Although there was a possibility that the organizations could be quite different, the participants' worked in essentially similar capacities. All participants needed to have worked on digitization projects to be able to speak about their experiences.

Relevant and meaningful results about shared perceptions or beliefs were expected upon speaking with four to six archivists. A saturation point could be expected to occur with twelve participants⁴¹. Data were used from the five people (in four institutions) that completed the interview and survey processes.

In order to answer the research questions, responses were needed from participants about how they've experienced the digitization selection process. Those sorts of responses required a qualitative approach rather than a purely quantitative one. Interviews were semi-structured and accompanied by a multiple choice and open-ended question survey for background information. The semi-structured format ensured some consistency across interviews, while enabling a degree of exploration. Because of the subjective (i.e. experiential and abstract) nature of the data, it

required an interactive discussion with experienced professionals. A semi-structured interview process ensured that specific topics could be covered but also recognized that different archivists would operate under different acquisition policies, scopes, etc.

Most of the interviews were recorded during face-to-face meetings. One interview was recorded during a phone conversation and another interview was conducted in text through e-mail. The audio interviews were transcribed and coded. The transcriptions were verified by an independent reviewer. The background surveys included seven questions and were all conducted via a one-page, Web form by the participant after he or she had completed the interview.

To analyze data, commonalities in participants' experiences were identified. Based on these commonalities, several recurring themes covering what archivists consider in their selection processes, surfaced, were coded, and used to assess the phenomenon.

Results

Participant Background

Four participants were female and one participant was male. Four participants were between the ages of 25 to 44, except one was in the 45 to 64 range.

All participants had worked in various capacities. The following list of work experience was most commonly repeated across participants, with everyone being involved in digitization project management and at least four people involved in each of the others.

- Digitization project management
- Cataloguing or description
- Participation in an advisory committee
- Preservation
- Selection of fonds to digitize

Other types of experience the participants commonly reported included accession work, nominating materials/fonds, nominations evaluation, and scanning. This range of experience

suggests a rounded understanding of digitization projects and not a focus on just one aspect of digitization. The following table provides an overview of the participants' backgrounds.

Participant	P1	P2	P3	P4	P5
Title	Director, Library Services	Digitization Librarian, Coordinator Digitization & Delivery	Director, Records Management and Archives	Acting University Archivist	Archivist at General State Archives of [Country]-Historical Archives of [Region]
Current role	Library, Director	Digitization librarian	Archivist, Director	Analyst, Archivist, Project manager	Archivist
Digitization projects worked on	2	4	20	5	4
Largest project resulted in # of digital objects	110	2500	100	20	7500000
Amount of work experience related to digitization selection.	2 – 5 years	Less than 2 years	More than 15 years	2 – 5 years	6 – 10 years

Table 1. Study Participants and Their Experience

Each of the participants had worked on 5 or fewer projects except for one participant that had worked on 20. The number of projects, digital objects created, or years of experience did not have an appreciable difference in attitudes or observations the participants made on selection nor the digitization process in general. It is worth noting that in spite of the way their roles are labeled, the participants were not always doing strictly what their labels would suggest. Namely, participant 1 [P1] also functioned as the institution's archivist.

There was a great difference in the size of the participants' largest projects. This indicates that the participants' responses to the interview questions reveal perspectives coming from a broad difference in terms of the magnitude of their work. Nevertheless, this difference did not reveal any obvious distinction in the participants' opinions.

Interview Topics

Topic A: What steps are taken in the digital preservation selection processes that archivists adopt in practice?

Most participants had not implemented a complete and formal policy for their processes such as those detailed in the literature. Two participants worked based on a variety of existing processes that were evaluated from project to project, but were planning to implement a formal system (e.g. Appendix A: P2, P4). Even in the absence of formal institutional policies, there were similar steps that the participants took. With the exception of one participant [P5], these steps were project-based: there was first some kind of recognition for a need to digitize a particular fonds or items within the holdings. That need was generally initiated by demand for the materials. This differs somewhat from what the participants would like to implement; to digitize based on institutional policy (ongoing and systematic) for handling their holdings and collections, rather than working project to project. In some cases, donors contributed funding for digitization, which essentially determined selection (barring prohibitive other factors, for example, copyrights).

The model, common among the participants, began with someone nominating a part of the collection for digitization. Nominations were not necessarily initiated through a predetermined procedure for digitization. The nomination could come from a request external to the institution, be initiated by the archivist or librarian, or else it could come from a colleague that wanted to use the resulting digital object. When the nomination was initiated by an archivist or librarian it could have been due to their recognition that the materials were becoming inaccessible (due to physical degradation or format obsolescence). In some cases, it was initiated in support of other types of initiatives their institution was undertaking or in recognition of what their user community wanted to access via the Web.

In P1's case, although the selection was expansive across the holdings, it was primarily to

ensure specific, continued access to degrading materials or obsolete formats. Still, we can consider this a demand-based project because in order to ensure the accessibility of the holdings, the archivist identified what of the holdings would be in demand and took steps specific for the purpose of her project, not based on institutional policy.

After the nomination for digitization, there were discussions on the nomination and the factors for selecting it; essentially an evaluation phase. Discussion could take various forms. In some cases it was a formal meeting among peers who were responsible for considering what was nominated and deciding whether or not to proceed. In other cases, the discussion was less formalized and operated as more of a collaborative review. These discussions served as a final check on the selection before submitting it to the next phase in the digitization process. If the discussions resulted in a negative decision, then the selection wouldn't be digitized. Typically, participants considered these factors:

- content (e.g.: significance, comprehensiveness, educational value or promotional value)
- the person or people generating the demand
- whether the digitized result would actually serve the intention
- physical condition
- the original format
- restrictions (e.g.: copyright or donor agreements)
- whether it was unique or not

Ultimately, the participants had repeatable steps they took to select for digitization. These were broadly similar to many of the existing guides in the literature even though they had not specifically adopted those guides.

Topic B: How do archivists perceive the influence their personal roles (including experience, attitudes, and skills) have in creating the intellectual object that results from a digital preservation selection process?

Participants recognized the personal knowledge they drew on and the objectives of their

position, role, and employer as influences on the selection process. Much of this knowledge came from their experience but it was also informed by the responsibilities of their profession.

Participants expressed that personal relationships lent them an informal understanding of researcher interests, what materials instructors would find useful, and other sorts of content appreciated by users (e.g. Appendix B: P1). These relationships and experiences influenced their approach to selecting, if not always directly then with a back-of-mind sense. The participants were self-critical about this influence—relying on discussion with their colleagues for a balanced perspective in selection (e.g. Appendix C: P3).

Beyond who the user was, the participants used their knowledge of what those users expected, to influence their selection judgments. They mentioned the assumption that people wanted access to everything (e.g. Appendix D: P2, P3), and they looked for insight about what people want digitized, examining things such as search criteria (e.g. Appendix D: P4). When participants recognized a user's expectations for access then they would approach the selection of certain materials with an eye toward what a person will be able to access once the materials are digitized.

As part of their job responsibilities, the participants were involved in activities such as the production of online exhibitions, building sites that enabled searchable indexes of digitized materials, made digital versions of existing finding aids, and set up thumbnail images for browsing and previewing. The participants recognized from user feedback or observation of current trends that if something was not visible through online searches and browsing, then users would get confused or believe that what is being sought does not exist (e.g. Appendix D: P4₂). Users may not think to go and physically visit an archive to find what they're looking for. This is one reason why a prevailing attitude was that *everything* ought to be digitized. Participants

recognized that they did not have the capacity to fulfill all user expectations but they gave consideration as to what would best serve their users.

Participants saw a need to increase awareness of their holdings and attract users. They used knowledge that they'd accumulated in their roles, to select what sorts of items would attract people to use the digitized materials (e.g. Appendix H: P3₂). They recognized that increasing Web traffic depends on another step: promoting awareness. They needed to do work to let people know that the digitized objects were available. P3 noted that measuring which social media posts garner more feedback, supported her decision to digitize more pictures. For example, good response rates from photo sharing service, Flickr, might influence her judgment on prioritizing images for digitization.

An intellectual familiarity with the content of their holdings influenced most participants' selection judgments and that familiarity takes time for archivists to develop (e.g. Appendix E: P1, P5). This appeared in several forms. In one sense it relates strongly to relationships with users because knowing what content will satisfy users' needs is crucial to serving them. Familiarity with the content also supported decision-making that had to contend with current affairs. P4 referred to her familiarity with certain materials related to WWI in support of researchers taking an interest due to the war's upcoming centenary. P3 mentioned the approaching anniversary of her institution, which was generating demand for documents of its past highlights to publicize. These examples show an awareness of subject matters, which the participants knew would be in demand. Participants felt the need to weigh the value of exposing this content against exposing other content for broader purposes, which was possible due to the familiarity they and their colleagues have of their holdings.

Participants expressed attitudes about how they might influence selection based on their

understanding of institutional objectives. Depending on the nature of the materials, participants deemed some more inline with institutional objectives. Or conversely, an archivist (e.g. Appendix F: P3) could, based on her experience and understanding of the institution's history judge that some materials would be potentially detrimental to the institution's reputation or painful for the memories of living people. While dutifully preserving these materials, the archivist would not deem them as much of a priority for digitization and thus the potential for broader dissemination. P2 reiterated the purpose of serving research needs, whether the needs were coming directly from members of the institution or even other institutions. Serving institutional objectives thus required participants to conceptualize and take to heart what the materials they selected to digitize would support.

The role of archivists in terms of how active they ought to be in the public sphere was influential to most participants' judgments. While they used their knowledge in supporting a role that contributes to events and current issues of interest, they did so in recognition of professional responsibilities. P4 spoke about another institution's example of a fonds, which included materials that the archivists considered might, if disseminated digitally to a broad potential audience, cause harm. "And so the ultimate decision was this information does not get disclosed... I would follow that because I think it can bring harm to the environment or to a certain situation and therefore... ethically and morally I don't think I could do that." This, exemplifies the sense that what archivists deem appropriate for digitization (and increased access) will impact the public sphere.

In sum, the participants identified how selections could be influenced because of knowledge they'd acquired from working with their users or colleagues. There was an influence on selection in the direction of digitizing everything (due to knowledge of users' search and

access expectations). This influence can be attributed to an understanding acquired in the course of their professional roles. Knowing the content of their holdings helped them serve institutional objectives, which thus brought some influence in line with selection. And they took account of their professional roles when considering ethical concerns for what the availability of a digitized object would mean to the public sphere.

Topic C: What are the recurring subjective factors in the identified processes that archivists make in their selections for digital preservation?

The participants described their selection processes such that the main phases can be grouped and summarized as nomination, discussion of nominations (evaluation), and the decision to digitize. These phases were influenced by participants' subjectivity through six common factors that relate to personal knowledge and professional perspectives. These factors were the archivist's

1. relationships with users and colleagues,
2. awareness of user expectations and needs,
3. intellectual grasp of the content,
4. understanding of institutional objectives,
5. stance on the extent to which his or her role should be active in the public sphere, and
6. identification of value.

The first five factors were exposed through the commentary on the influence of personal roles in Topic B but identification of value was expressed in other ways.

Participants were concerned with value from an intellectual perspective; an issue for researchers. Participants (e.g. Appendix I: P1, P5) talked about the value of the archive's holdings in terms of its highlights. Determining what makes something a highlight of a collection was something that the participants intuited through their professional background or professional awareness. They presented arguments and examples to justify highlights. P1 offered examples

through her knowledge of her community's ongoing recognition that a particular conference had been a hallmark for the arts community (the archive had tapes of its proceedings) or from poetry that she understood had acquired a certain renown.

Another way of assessing value was based on a knowledge of completeness. P4 relied on her professional knowledge of how much of a fonds existed within her institution's holdings. If she knew that significant portions existed at other institutions or simply were missing from her holdings, then she might not judge what she had as valuable for digitization. She stated

Once we've decided on a fonds and then we would go through and actually make sure that... it's complete enough that it will actually produce a story—there's, I find sometimes that you may only get one or two items in from the donor, or it could be a one-off, so you may not necessarily want to digitize those as much because there's not enough context or provenance around it in order to make it a story to put up.

She wanted to be able to provide a full “story” for users to access and if she understood that she could only provide users with a fragment of what was potentially available, she judged that fragment to be less valuable for digitization.

Evidential value to the institution concerned participants P3 and P4 who also held records management responsibilities. P3 spoke of evidential value, documents serving the university, and informational value. By way of explanation, P3 offered a hypothetical example that if there were records documenting the Beatles performing at one of the institution's facilities, those would have high informational value because of their significance to the overall community even though this record of the institution's history had little impact on its objectives. Recognizing this requires the archivist to have a personal awareness of the cultural and historical significance of such an event and to consider the relevance of that within the context of her function.

Participants also considered value in terms of what materials could attract more users to the archives. For example, materials that could enable collaboration with other institutions for

exhibits that would attract attention might be considered more valuable to digitize.

Digitization guides instruct using commonly recurring factors such as fragility or frequency-of-use in order to nominate and select materials. Comparing those types of factors to the six listed above distinguishes what informs an archivist's subjectivity. Rather than being based on objectively identifiable physical characteristics or statistics, these six repeating factors stem from knowledge or perspectives that the participants themselves exhibited. These are the entry point to examining an archivist's subjectivity during decision making processes.

Discussion and Recommendations

The subjective judgments that archivists make are a key part of selecting materials for digitization. These judgments are based largely on archivists' personal knowledge and professional perspectives. We ought to be concerned with these judgments because an individual's subjectivity influences selection in a way that is neither transparent to those who will use the digitized objects, nor to the people that will maintain the digitized objects into the future.

We looked at the steps taken in the digitization selection processes that archivists practice. Can these steps, documented in the existing guidance, be extended to include uniform procedures for dealing with subjectivity? Due to the idiosyncratic nature of subjectivity, it may not be possible to make a perfectly comprehensive guide on how to approach every subjective factor in a selection phase. Evidence of this is plentiful in the literature, which recognizes the subjectivity that exists as part of selection decisions but does not offer a comprehensive step-by-step procedure to approach it (at least not with the granularity it does for objective factors).

We could ameliorate this by providing greater transparency about the way subjectivity factors into the selection. This would help users interpret and understand the context of what they access. This proposition is to formally record subjective judgments as a set of selection notes,

which correspond to the digital object. This could be done without requiring a great amount of detail. To look at the rationale for this and how it might work, let's consider the following items. First, why we digitize (looking at what is practiced), the result of digitization, and the influence of the professional role with respect to the expectations and responsibilities around the digitized object. This last point is informed by the subjective factors that recur.

Why We Digitize

It's important to reflect back on the point that digitization is done primarily to enable access—a goal that is also true when the need for preservation motivates digitization. Digitization does not preserve the original but it aids preservation. The archivist gains more ability to limit handling of the original and even if the original becomes totally unusable, at least a digital version can be carried into the future. The participants in this study expressed perspectives that were consistent with the literature on this access/preservation goal. It satisfies some essential user expectations such as P1's statement on the deteriorating materials in her collection “Over the years we knew that these materials were becoming more fragile and so we just weren't letting people use them very much for research or teaching purposes.” In a situation where the original analog materials are too fragile for people to continue accessing, then accessing a digitized version offsets the wear that would otherwise be put on the analog original.

The problem is that the participants felt that users *assume* they'll have digital access to everything. A simple view on meeting users' expectations then dictates that everything should be digitized. If that is the case, then why follow approaches in the literature for selecting at all? Archivists must enter into a selection process as a contingency of resource constraints (time, money, personnel, equipment). We can see this reflected in comments such as P1's about how long it will take to accomplish her goals “I know that these things can take time especially when

you have no money, no staff... there's so little funding for archives..." Since it is not possible to digitize everything, these constraints entail selectivity—a process that relies on factors and judgments documented in the literature and this study. The constraints largely revolve around funding, personnel, equipment, time (to do the digitization work), and creation of an additional, ongoing maintenance responsibility. So long as the existing constraints prevent digitizing everything, archivists will need to select that which best serves the requirements of access.

The Result of Digitization

The act of digitizing, results in the creation of a new object, which did not exist before. The responsibility of enabling access to this new object calls for an understanding of the selection process. Although we may label it a digital surrogate or a digital version of the original (analog object), this new object has very different characteristics from the analog. We ought not think of digitization as merely a surrogate or as a copy of an object. It's better to understand that the new object does not convey the original's physical characteristics, adjacency to other objects or physical context, etc. It is its own object retaining an intellectual link to the physical object from which it was initially created.

Existing literature distinguishes the new object when detailing the digitization process. Guides frequently use terminology to explain the process as “making” or “creating” the digital object. A guide such as the Jones and Beagrie handbook even devotes a section to how digital materials are different⁴² from the analog sources. Indeed, archivists distinguish between the digital object and its analog source; typically denying that the digitized object can serve as a replacement for the analog source. Rather they preserve both. If one object was essentially the same as the other it would be difficult to argue that a creative act is involved; but then archivists also wouldn't need to preserve both. Given the resources and effort required for preserving these

objects, coupled with the often difficult fiscal position of archival institutions, this distinction bears a significant weight. Those who create and maintain this new object, face the fact that they've made it possible to gain access to something, which people either could not previously access or can now access in a new way.

Influence, Expectations, and Responsibilities

When they select and digitize content to make it accessible to their users, archivists are creating something intended for improving long-term accessibility. Professional practices and policies are acutely important to the role so we ought to characterize the digitized object to gain a sense of what's expected out of the archivist's role and accompanying responsibility.

A useful characterization of the digitized object comes from Paul Conway who explains that we are not dealing with preserving an object's physical integrity but “specifying the creation and maintenance of the object whose intellectual integrity is its primary characteristic.”⁴³ Terry Cook offers insight about the role, saying that the archivist is becoming “...an active mediator in shaping collective memory through archives. Archivists inevitably will inject their own values into all such research and activities, and thus will need to examine very consciously their choices in the archive-creating and memory-formation process.”⁴⁴ Though writing about appraisal, the point is useful in selection for digitization⁴⁵.

Just as selectivity in the appraisal process for the original archival materials promotes certain characteristics in the holdings overall, so the selectivity of the digitization process influences the character of what users will receive intellectually from these new digital objects. The resulting impact of these processes can be quite important. For example, Verne Harris looked at the tendency to think of archives as reflections of reality in which the archivist is essentially neutral. Harris disputed this tendency with a variety of examples about South African archives in

relation to the apartheid system. Harris says

The appraiser's values, quality of work, perspectives, interaction with the records creating agency, engagement with the policy s(he) is implementing, and so on, all shape and are reflected in the appraisal. The appraiser is not simply identifying records with archival value; s(he) is creating archival value... So the archival record provides a window into the appraisal process as much as it does into anything else. This demands an approach which embraces the individual appraisal report or study itself as a record, a text with a very specific ontological status.⁴⁶

Likewise, in the context of digitization the archivist is creating something that upon being accessed will carry some implicit quantity of information about the process and judgments that resulted in its creation. That is, information about the individual archivist's subjectivity.

Archivists, as we see in the literature, traditionally adopted a position of impartiality. Impartiality can be supported by unambiguous, objective factors for selection so that a decision can be evaluated from a neutral position independent of a particular person's own experience. However this has to be squared against the fact that participants in this study also used their personal knowledge in selection judgments.

In fact, the experiences of the interview participants show that they have a responsibility for an active personal influence in their roles. While recognizing that there is an onus to provide access to digitized materials, they noted examples where their judgment could be actively called upon by senior institutional members. P4, for example, explained how the principal or other senior official of the institution could make requests for materials that support an important upcoming event, relying on her judgment to determine what would be relevant and useful.

The participants acted with some responsibility based on the background awareness they had of what is occurring in the public sphere. This awareness also extended to things such as judging potential benefits or harms that could ensue from digitized materials. Thus, selection for digitization does not come from a completely impartial stance and what is created, like Harris's

commentary on the ontological status of appraisal, has its own distinct existence.

Addressing the Impact of Subjectivity

How can we address the impact of subjectivity in selection, if not with a step-by-step guide for archivists to adhere to? We can increase selection transparency: record notes about the process. These would be akin to scope and content notes, which provide the people that access archives with information and context for what they will find in a fonds. In the case of the digitized object, the context for its existence becomes more complete with information about why the particular digital object was made. Digitization notes could serve the purpose of letting an archivist reveal what factored into his or her judgments.

This task would add to the digitization effort, which is a worry considering the already strained availability of resources in institutions. But such notes would divulge the influence of subjectivity so that others could gain awareness and consider the factors that influenced what they're accessing when using or maintaining the archives. In light of the impact that the digitized objects have—how they shape context and perceptions, increasing selection transparency is worth consideration.

Documentation for selection transparency could begin by noting the information or rationale behind each of the six recurring factors that related to archivists' personal knowledge and professional perspectives (there may be others to include).

1. relationships with users and colleagues,
2. awareness of user expectations and needs,
3. intellectual grasp of the content,
4. understanding of institutional objectives,
5. stance on the extent to which his or her role should be active in the public sphere, and
6. identification of value.

The notes might be simple. For example, number 2 might have a note such as “Known to be used

in a course on surrealist games.” In addition, while these six factors tend to recur, questions involving subjectivity that are raised in existing guides prompt useful notes. The result of this reflection is a record that illuminates the existence of the selected object (similar to Verne Harris's description of the window on the appraisal process).

Let's look at some reasons why archivists' personal knowledge and professional perspectives would contribute to users' understanding of the context for what he or she accesses.

Personal Knowledge

As we've seen, archivists' personal knowledge of user expectations, content, and institutional objectives influences selection, which we place within the context of an effort to make what is digitized, accessible. Selecting to digitize something is a choice to digitize *for the thing's users* so archivists have to think about what users expect in terms of access to the digitized object. The knowledge that archivists accumulate throughout their experience working with their community is paramount here. The fact that archivists use this knowledge was evident in the participants' responses on the theme of user expectations—comments regarding users, such as P1's “You get to know what everybody's teaching and what everybody's research is framed around. You just know that they're your people and they're the ones that you keep in mind all the time...” indicate ongoing thinking about the people accessing the archives but those thoughts are hardly transparent as P2 expressed “...he was so annoyed at the fact that we kept on doing this digitization where we'd only show a partial of it... For him, what it meant was that he'd find five letters but he'd have no idea if that was all of the letters, if that was all the letters that were digitized, or if that was just the letters that had been selected.” We must try to understand the assumptions that users will make in perceiving what they find. This effort presents itself in

different forms.

It's unlikely that an archivist could anticipate every possible way that a person might perceive what they access because the uses to which people put digital materials are wide and varied, even in a purely research context. Digitized objects enable access in ways that were not possible for the original. Since digital objects do not have spatial requirements, they can be accessed from any place and a variety of devices, outside the context of the whole physical extent of the fonds. Due to their copyable nature, their accessibility leads to being incorporated in new ways, thus accessible from within other objects that have nothing to do with the original. There are popular examples of this through music sampling or artwork⁴⁷.

If however, the user is given some information about the way in which the archivist's role influenced the selection, he could actively interpret what he is accessing. Hypothetically, the person in P2's example of digitized letters may not be so annoyed because beyond knowing the extent of what was digitized, he could understand that only some letters were chosen for digitization: the letters that were the result of a project the institution had for an event that it wanted to highlight. This would give him the sense that the archivist had determined these were especially good exemplars for the topic of that particular event. If his research was on a subject different from the event topic, he might be more inclined to seek out the other letters.

Professional Perspectives

Identifying the value of the materials to digitize relies significantly on the archivist's professional perspectives. Literature such as Hazen, Horrel, and Merrill-Oldham's work, which offered ways to judge intellectual value⁴⁸, sets the problem of value before archivists and indeed the participants in this study made value judgments in the scope of their professional practices.

While users have their own reasons for valuing some particular content that has been digitized, their reasons may not be identical to the ones archivists used for their selection. Communicating the value judgment that went into the selection would help build context for the user to better understand what he or she accesses digitally.

Let's refer back to P3's example of the Beatles performing at one of her institution's facilities. Someone accessing the digitized records might be evaluating the university's history of use of its venues for cultural events. If the records had been selected due to their value documenting events at the university's venues (not because it was the Beatles performing) then the user might gain a more immediate inkling that such a record is an example available within other like records and thus perceive a more relevant context for his research about venue usage. However, if the documents were selected for digitization due to the influence of their significant informational value (a performance by the Beatles) that would express a particular background for why they're digitally accessible: star power—it does not suggest that the record is available within a context covering the usage of university venues.

Among participants in this study, value (usually involving information content and significance) was judged according to

- highlights of the institution (requires knowledge of the institution's objectives)
- items held in esteem by the user or potential user community (requires knowledge of users)
- ideas such as completeness of a fonds for digitization (requires knowledge of the content)

These areas do not adhere to well-defined objective factors, instead they involve the archivist's professional perspectives. Such judgments do require some objective data but due to their intellectual and experiential nature, they also involve the consideration of a lot of grey areas to make a justifiable selection.

Highlights make for a slippery criterion. In referring to P1's example of highlights, she had a sense of what the communities interested in certain materials considered important. She identified things that were important to many people within her user community but that she herself didn't care for (e.g. the poetry). Her knowledge in this case came primarily from her regular interactions with professors (people in the university using the archival materials). While that value is currently understood within her community, the archive will maintain the materials for a long time and that understanding may not always be a given. Documenting her professional perspective of what the community values, adds context for the selection.

Hazen, Horrel, and Merrill-Oldham's assessment questions cover whether value warrants the access that would become available because of the digitization project⁴⁹. Recognizing the subjectivity involved in that assessment, Hazen, et al., offer ideas on how to judge it. First, they suggest that researchers would generally prefer electronic versions of the originals as opposed to derivative works. They'd place a higher intellectual value for digital access than access to a derivative. Prominent scholar's papers would also be given a higher intellectual value than administrative papers. Next, they ask for some consideration as to how digitized copies might be assembled in aggregate to increase the value of what becomes accessible. They point out that often times researchers can get more value from things when they see them in aggregate. Essentially, if archivists select materials of the sort of content that can be linked and collocated with other collections, then the digitized materials will provide a greater context and more comprehensive collection, and thus presumably represent greater intellectual value. On the other hand, studying usage patterns of the materials may show that some are rarely used and that may be reason to consider them of lower intellectual value. Hazen, et al. admit this is a problematic way of judging because there are many reasons aside from intellectual value that could contribute

to low use. These ways of identifying value can raise useful awareness during a selection process but they rely on the archivist's professional perspectives. The practice also means that archivists must presuppose what will be of value to a researcher, which is a difficult task, considering the value of the materials will vary in unanticipated ways and according to the researcher's objectives.

In all of those methods for judging value, the archivist's professional perspective on how the value of the fonds was identified could be recorded. A record of those judgments, promotes transparency that would enable the user to validate his or her interpretation of what is being accessed with the value the archivist identified for the digitized object.

We need to also consider the archivist's professional activity in the public sphere. One lens on interacting with the public sphere comes in the form of selecting materials to digitize based on how the end result can be linked with other collections. Archives taking an active approach in presenting their digitized content are necessarily implicating themselves in public life, in such a way that they make choices, which reflect certain subjective judgments. Tom Nesmith offers the example of how archives, anticipating the way people want to access digital information, select content that helps to provide context for popular subject matter. "The Canadian Broadcasting Corporation (CBC) makes available (at both its English- and French-language services' websites) television and radio clips from its archives that offer insight into the historical context of many current issues, events, and the public figures associated with them. As many as twenty clips may accompany a particular subject.⁵⁰" Digitizing objects to be linked with certain collections or resources as opposed to others is a judgment on the sort of research or interest that will be associated with them. This puts the archivist into a role where he or she is creating something prepurposed for a topic, as opposed to enabling its use from an impartial

perspective.

Finally, we ought to reflect on the way participants drew on knowledge they'd acquired through their relationships. In educational institutions, for example as P1 explained, her knowledge of the types of materials teachers liked to use for their courses was important to keep in mind. P2 talked about being more willing to undertake digitization projects when the project included people that had been involved in previous projects. We cannot expect ordinary users of the digitized materials to have any awareness of the archivist's professional relationship or how that would influence what is accessible. But a note identifying that something had been digitized in the course of repeated work with internal stakeholders could help a user to understand that what they're accessing may have been the result of some kind of institutional access requirement. This for example, might lead to questioning the availability of other resources, which might be available but did not need to be digitized.

Conclusion, Limitations, and Further Study

In this study, I aimed to present research and recommendations that help deal with the impact of subjectivity and complement the existing guidance. A way to give more context about selection would be through digitization notes based on the six recurring factors of subjectivity. Such a practice would make explicit, how archivists relied on their personal knowledge and professional perspectives. Transparently documenting what the archivist had in mind where subjectivity enters the selection process could do much to help the user perceive and understand what he or she accesses in a digital context.

One of the limitations of this study is that the majority of the participants worked as part of an academic institution—just one worked in a government context. This fact limits the variety

of projects that the participants spoke about and may be somewhat restrictive in terms of their perspectives on organizational processes and peers. Perspectives from cultural, local civic organizations, or business entities could expand the understanding of archivists' digitization experiences. It is possible that taking time to interview more people would yield additional perspectives from a broader range of experience.

Further study that focuses on users' impressions about how they perceive the context of what is available to them when accessing digitized objects, would be valuable. Rather than relying on search and access statistics or archivists' anecdotal evidence, such research could help characterize users' interpretations. It would enable further development of a way to divulge necessary information about which objects were selected from within the holdings or collections. It would also be useful to research and test different methods or formats of divulging subjective selection information to determine what would be most effective for users.

Finally, while not adhering to a specific, existing set of digitization guidelines, participants did tend to practice similar steps during the selection phase: nomination (using common, largely objective factors), discussion or evaluation of what had been nominated, and a decision on whether to digitize the nominated material or not. Participants recognized that the experiences, attitudes, and skills they'd developed, influenced their selection judgment. These ranged from their personal familiarity with users and knowledge of users' expectations, to knowledge of the content of their holdings, to obligations toward their institution's objectives, and professional perspectives on the nature of their activity within the public sphere. In the course of this study it was possible to identify recurring selection factors, which injected subjectivity into selection.

The participants' experiences did not contradict the literature, which suggests that

archivists' selections can shape understanding and discourse of a subject. In fact, this shaping is what makes the concern over subjectivity important—it is a departure from the impression or expectation of neutrality and we ought to account for that on a professional level. When the participants undertook digitization processes that were similar to those detailed by the guides, their judgments were not wholly determined through objective factors but also influenced by their subjectivity. It may not be possible (nor even desirable) for archivists to try to select solely based on objective factors but they could consider ways to give people a more informed context, allowing better understanding of what users access digitally.

References

- Anderson, Cokie. *Ethical Decision Making for Libraries*. Oxford, UK; Chandos Publishing Ltd., 2006.
- Association for Library Collections & Technical Services, Preservation and Reformatting Section. 2007. "Definitions of Digital Preservation". ALA Annual Conference Washington, D.C., June 24, 2007.
- Astle, Peter J, and Adrienne Muir. 2002. "Digitization and Preservation in Public Libraries and Archives". *Journal of Librarianship and Information Science*. 34, no. 2: 67-80.
- Avery, Cheryl, and Mona Holmlund. 2010. *Better off forgetting?: essays on archives, public policy, and collective memory*. Toronto: University of Toronto Press.
- Ayris, Paul. "Guidance for selecting materials for digitisation." (1998). Accessed from http://eprints.ucl.ac.uk/492/1/paul_ayris3.pdf
- Baillargeon, Stéphane. "La mémoire de l'art." *Le Devoir*. October 8, 2014. <http://www.ledevoir.com/societe/medias/423285/la-memoire-de-l-art>. (accessed October 9, 2014).
- Bradley, Kevin. 2007. "Defining Digital Sustainability." *Library Trends* 56, no. 1: 148-163. *Library Literature & Information Science Full Text* (H.W. Wilson), EBSCOhost.
- Bülow, Anna, Jess Ahmon, and Ross Spencer. 2011. *Preparing collections for digitization*. London: Facet Publishing, in association with the National Archives.
- Canadian Council of Archives. 2002. "Digitization and Archives". Accessed from http://www.cdncouncilarchives.ca/digitization_en.pdf
- Canadian Council of Archives. 2008. "RAD Appendix D Glossary" Accessed from www.cdncouncilarchives.ca/RAD/RAD_Glossary_July2008.pdf
- Canadian Heritage Information Network. "A Guide for Managers Planning and Implementing Digitization Projects" rcip-chin.gc.ca. 2000. Canadian Heritage Information Network. Accessed on http://www.pro.rcip-chin.gc.ca/contenu_numerique-digital_content/numerisez_collections-capture_collections/publication-publication/numerisez-capture-eng.jsp (2014)
- Carignan, Yvonne. 2005. *Who wants yesterday's papers?: essays on the research value of printed materials in the digital age*. Lanham, Md: Scarecrow Press.
- Conway, Paul. 2010. "Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas." *Library Quarterly* 80, no. 1: 61-79. *Library*

- Literature & Information Science Full Text (H.W. Wilson), EBSCOhost.
- Cook, Terry. 2001. "Archival Science and Postmodernism: New Formulations for Old Concepts". *Archival Science*. 1 (1): 3-24.
- Fresko, Marc, and Tombs, Kenneth. 1998. "Digital Preservation Guidelines: the State of the Art in Libraries, Museums and Archives." European Commission. Directorate General XIII-E4. pp 41 – 47. Accessed from <ftp://ftp.cordis.lu/pub/ist/docs/digicult/study1.doc>
- Gertz, Janet. "Selection for Preservation in the Digital Age." *Library Resources & Technical Services* 44, no. 2 (2000): 97-104.
- Guest, G., Bunce, A., and Johnson, L. 2006. "How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability". *Field Methods*. 18 (1): 59-82.
- Hampshire, Edward, and Johnson Valerie. 2009. "The Digital World and the Future of Historical Research 1". *Twentieth Century British History*. 20, no. 3: 396-414.
- Harris, Verne. 2002. "The Archival Sliver: Power, Memory, and Archives in South Africa". *Archival Science*. 2, no. 1-2: 1-2.
- Hazen, D., Horrel, J., and Merrill-Oldham, J. 1998. "Selecting Research Collections for Digitization." [clir.org](http://www.clir.org). Council on Library and Information. <http://www.clir.org/pubs/reports/hazen/pub74.html#intellectual>
- Hughes, Lorna M. 2004. *Digitizing Collections: Strategic Issues for the Information Manager*. London: Facet.
- Jenkinson, Hilary. 1922. *A Manual of Archive Administration Including the Problems of War Archives and Archive Making*. Oxford: The Clarendon Press.
- Jones, Maggie, and Neil Beagrie. 2001. *Preservation Management of Digital Materials: a Handbook*. London: The British Library for Resource, the Council for Museums, Archives and Libraries.
- Lor, Peter Johan, and J.J. Britz. 2012. "An Ethical Perspective on Political-Economic Issues in the Long-Term Preservation of Digital Heritage." *Journal of the American Society for Information Science and Technology*. 63 (11): 2153-2164.
- McIlwaine, John, Jean-Marc Comment, Clemens de Wolf, Dale Peters, Borje Justrell, Marie-Therese Varlamoff, and Sjoerd Koopman. 2002. "Guidelines for Digitization Projects for Collections and Holdings in the Public Domain, Particularly Those Held by Libraries and Archives" (March): 59. <http://archive.ifla.org/VII/s19/pubs/digit-guide.pdf>.
- Moss, M. 2005. "The NINCH Guide to Good Practice in the Digital Representation and

- Management of Cultural Heritage Materials Technical Guidelines for Digital Cultural Content Creation Programmes." *American Archivist*. 68, no. 1: 178-181.
- Pymm, Bob. "Building Collections for All Time: The Issue of Significance." *Australian Academic & Research Libraries* 37, no. 1 (March 2006): 61-73. *Library Literature & Information Science Full Text* (H.W. Wilson), EBSCOhost.
- Schwartz, Joan M., and Terry Cook. 2002. "Archives, Records, and Power: The Making of Modern Memory." *Archival Science*. 2 (1-2): 1-2.
- Sitts, Maxine K. 2000. *Handbook for Digital Projects: a Management Tool for Preservation and Access*. Andover, Mass: Northeast Document Conservation Center.
- Smith, Abby. "Why Digitize?" *clir.org*. 1999. Council on Library and Information Resources. Accessed on <http://www.clir.org/pubs/reports/pub80-smith/pub80.html#access> (2014)
- Sternfeld, Joshua. 2011. "Archival Theory and Digital Historiography: Selection, Search, and Metadata As Archival Processes for Assessing Historical Contextualization." *American Archivist*. 74, no. 2: 544-575.
- Still Image Working Group. "Digitization activities project planning and management outline. US National Archives and Records Administration." Accessed May 13 (2014): 2010.
- United Kingdom. The National Archives. *Corporate Preservation Policy*. [Richmond], 2009. <http://www.nationalarchives.gov.uk/documents/tna-corporate-preservation-policy-2009-website-version.pdf>
- United Kingdom. Parliamentary Archives. 2008. "Digitisation Guidelines". Accessed from www.parliament.uk/documents/upload/digitisation-guidelines.pdf
- United Kingdom. Parliamentary Archives. 2009. "A Digital Preservation Policy for Parliament". Accessed from <http://www.parliament.uk/documents/upload/digitalpreservationpolicy1.0.pdf>
- Zhang, Allison B., and Don Gourley. 2008. *Creating digital collections: a practical guide*. Oxford: Chandos Pub.

Appendices of Participant Commentary

The selected excerpts of commentary provide supporting examples for the issues discussed in this study.

Appendix A: Formal Processes

Participant	Commentary
P1	“...then we had to start thinking, we meaning me, I had to start thinking about preservation so I started not even systematically but on a sort of a case by case basis prioritizing which videos I want to migrate to a digital format. I did that in a week by week basis when I decided that we need to really start preserving a lot of these items. The way that I determined that was based on the content...”
P2	“Our current model is project based. So somebody nominates a collection or a project, it goes through a form that we have online and basically there needs to be a sponsoring librarian but it could be anybody. That puts forward the project, they fill out the form, it goes to a digital projects committee made up of various people throughout special collections and archives and they review it based on criteria that is also on the web site for collection guidelines. Then we basically say yeah or nay.”
P3	“We don't have anything formal for now. We would like to but it's something we are developing... What we would like to do is do the processing of one fonds and then in the end, digitize it all, all of the material there... It's not perfect and when there's an external researcher asking for something, then we digitize as well. Sometimes it's also based on preservation needs.”
P4	“If we have someone who would like to have something digitized, so give a donation to support the project... Then it's sort of already been preselected or predetermined based on what the donor wants to have digitized.”
P4 ₂	“Right now, they're general best practices but they are going to be made into a guideline... So we're actually working towards something a lot more substantial and more objective but for right now, it's quite subjective in terms of how we're doing it, it's based on our impressions in how we're deciding it”
P5	“The digitised archival material came from 37 agencies all over [the country]. Given that only a small portion of the collections could be digitised, the proposed material from each agency had to meet (as much as possible of) the following—and intercrossing sometimes—main criteria: multiple archival value; very frequent use in the reading room and/or requests for copies; physical condition of the original restricting access; material best representing the collections; highlights of an important archives; educational value; unknown (to the research) material to be promoted.”

Appendix B: What Archivists Know Due to Personal Relationships

Participant	Commentary
P1	“You get to know what everybody's teaching and what everybody's research is framed around.”

Participant	Commentary
	You just know that they're your people and they're the ones that you keep in mind all the time..."
P2	[<i>On digitizing via a project-by-project suggestion basis as opposed to an overall policy and strategy</i>] "...putting forward project-based digitization, I would argue, is not very good. Because it is self-selecting. I mean self-selecting isn't always the worst thing, these are smart people that are self-selecting that are doing it for the greater good that are looking at their researchers, that are talking to the researchers but it's still by nature, self-selecting and it's exclusionary. That's what I worry about, that it's exclusionary."
P2 ₂	"...you're more likely to go forward with a project with people that you've worked with in the past. That totally influences—if they've done projects with us before, they know the drill, it's easier."
P2 ₃	"My own personal knowledge has influenced it a lot... I worked in special libraries, so I worked with very different user sets. So that's dramatically influenced... I did a lot of visual research in other people's archives and I did all the acquisitions for visual images from other people's archives, and then I also worked in an archive where I sold images to other people. I worked a lot with media companies and film companies. That heavily influences me, in terms of what I think people want. How people use it, how people might search for it because I did so much of it before."
P3	"We are here to serve two people. The... employees, including teachers, faculty, and staff; and our students and all the external researchers. When we digitize a fonds it's mostly for employees because we want to create a sense of belonging... while disseminating archives... when there's an external researcher asking for something, then we digitize as well."
P4	"...we've made exceptions and stuff like that with photographs because people love photographs. It's nice... people want to scan it to use it as a gift, use it in a publication, whatever."
P5	"Familiarity with user needs is essential. For a coordinator, an experienced team on these matters is crucial. Inevitably, personal choices in every stage of the process shape the final product."

Appendix C: Checking the Influence of Personal Knowledge

Participant	Commentary
P2	[<i>On the people involved in digitization decision-making groups</i>] "...everyone should have a voice but as soon as you start getting larger and larger groups it starts getting more and more difficult to get things done and so we're trying to find that sweet spot of making sure that different collections are represented and different view points are represented"
P3	"...a background helps us to do a selection or something of interest to one of the archivists. And I think it's something that will happen in every institution unless we have very formal and not emotional—but to me, archives are emotional... I trust my team. They know better than I do... And my personal interest, I don't know, I try not to put them out there, not to put them out first but sometimes I may do it without knowing it. My team would tell you."
P4	"We also make the decisions... as a committee. It's not done in isolation."

Appendix D: What Archivists Know of Users' Expectations

Participant	Commentary
P2	“The assumption that we could select the best thing that they wanted to see was the assumption before, the assumption now—and the research is really showing now that no, people just want access to everything.”
P2 ₂	[<i>Speaking about a particular researcher</i>] “...he was so annoyed at the fact that we kept on doing this digitization where we'd only show a partial of it... For him, what it meant was that he'd find five letters but he'd have no idea if that was all of the letters, if that was all the letters that were digitized, or if that was just the letters that had been selected.”
P3	[<i>Speaking about the next generation of users</i>] “I think as we go, if you want to survive as an archive you have to move to be on the web... They will think that everything is digitized.”
P4	“The finding aids are really how people can find the material at a high level, and then from there, we can direct them more to a lower level, item level selection... Most of the stuff that we have done are more like online exhibits and here's a collection, here you go... I think our first step is going to be getting the material, the finding aids, into catalogues so people know what is there and then from there we'll look at how they're searching and what they're doing for search criteria to see how we might do the digitization, the metadata, and stuff like that.”
P4 ₂	“I think for most people, if it's not on the Web, it doesn't exist. So if we don't have it up there, they don't know we have it. I think their expectation is to have everything up and have it somehow or other searchable... I think people want it all, they want it now, and it has to come up, really high up in their hit list.”

Appendix E: What Archivists Know About the Content

Participant	Commentary
P1	“I don't think someone who just came in would have the same kind of instincts about what to preserve but there would be people around who could advise someone... but personally it was my own knowledge. I think that was important in prioritizing those items.”
P4	“...sometimes it boils right down to your gut instinct, your gut intuition... it has to actually do with knowing your institution and what you have in your institution.”
P4 ₂	“...you may take a special interest in it yourselves, that might drive something. Sometimes that's how exhibits get chosen, if there are special interests in a particular area, there may not be a great drive from outside, from external, but because we're taking an interest in it and we want to highlight something, we may digitize it and put it on as an online exhibit.”
P4 ₃	“I like to be a bit cautious about it being too researcher driven. Because right now for the next four years there's going to be a lot, at least for last year and this year probably next year for sure, there's going to be a lot of drive on WWI and the history of WWI because of the centenary. But once that's over, the drive isn't going to be there so if we make this big mad rush just to digitize anything related to that, we're ignoring something that can be more long term or can have more benefit or use.”
P5	“...personal knowledge of the collection's content and physical state also.”

Appendix F: What Archivists Know About Their Institution's Objectives

Participant	Commentary
P3	“...we need to have a good knowledge of the history of the institution...”
P3 ₂	“And we need to control the image of [the university].”
P3 ₃	[<i>On a tragedy in the institution's history</i>] “...we have a very famous event that happened twenty years ago, where... killed his colleague but we don't want to advertise that because it's not something that is—of course it's something that we want to remember the victims but we don't want to put him as a hero or something. So we try not to advertise anything related..”
P4	[<i>On digitizing materials for the sake of institution-specific events</i>] “Sometimes it's directed through a senior executive so it could be the principal, someone else, the provost, vice provost, someone like that. He could be saying, ok this is coming up, we need you to become more involved, we would like you to do something on this.”

Appendix G: The Active Role

Participant	Commentary
P2	“I think we have a massive responsibility to the public sphere, just from a collecting standpoint. We collect our histories. We have that responsibility in that sense.”
P3	“I think we need to have a good knowledge... of what's going on within [the institution] or within the scope of [the province], and maybe international... it may be there in the background but it's not something formal that we say, oh I noticed this movement or this way of doing...”
P4	“We're definitely becoming more active in these things... Sometimes it's directed through a senior executive so it could be the principal, someone else, the provost, vice provost, someone like that. He could be saying, 'ok this is coming up, we need you to become more involved...”
P5	“I think this is one of the goals of a digitisation project: provide access to information that matters, for a variety of reasons.”
P5 ₂	“The final goal of keeping archives is to make the contained information available to the public (Code of ethics and Universal declaration on archives of the ICA describe best the duties and commitments of an archivist...)”

Appendix H: Attracting Users

Participant	Commentary
P1	“You know I realized the digitization was the most important thing to do at that time and then what do you do with the content after that? We want to make it available and we want to also use it as, you know, leveraging our collections to attract more.”
P3	“If we don't let our client know it's there, they won't even search for it. They want to have everything there, they're so used to it with Twitter, with Facebook... If you don't put it on your Facebook and it's only on your Web site then no one will know it exists. So sometimes we choose to advertise digital material...”

Participant	Commentary
P3 ₂	“For photos we try to digitize the entire photos but we go with what people want or what has an impact. We can calculate or measure that impact because we have Google analytics on the Web site so we can measure what works well or what doesn't work well. Also on Flickr we can see how many hits... on my Twitter account I can also see who's following me, and ... what is the most viewed—that creates feedback. That's how we measure. Yes, it has an impact, that's why we are digitizing more pictures than anything else because that's what people want to see.”

Appendix I: Value Judgments

Participant	Commentary
P1	“If there's something that's really unique to our collection and very much about our institution then that would be a highlight.”
P3	“...what is important is more the evidential value, what documents serve the university, then informational value as well.”
P4	“Once we've decided on a fonds and then we would go through and actually make sure that... it's complete enough that it will actually produce a story—there's, I find sometimes that you may only get one or two items in from the donor, or it could be a one-off, so you may not necessarily want to digitize those as much because there's not enough context or provenance around it in order to make it a story to put up.”
P5	“...material best representing the collections; highlights of an important archives; educational value.”

Appendix J: Survey Instruments

Part 1 – Background Information

Please respond to the following survey questions. These will be used to understand a background context for your answers during the interview. Your identity will remain confidential.

- **I hold a job function as:**

(select all that apply to your position)

- Analyst
- Archivist
- Conservator
- Preservation librarian
- Developer
- Project manager
- Rare books librarian
- Other (please specify)

- **I have the following types of work experiences.**

(select all that required your involvement in digital preservation selection processes)

- Accession
- Cataloguing or description

- Retention
- Conservation
- Preservation
- Development (software, technical, or design)
- Data entry
- Document preparation
- Nominations evaluation
- Digitization project management
- Nominating materials / fonds
- Participation in an advisory committee
- Quality control
- Scanning
- Selection of fonds to digitize
- Other (please specify)

• **I have this amount of work experience related to digital preservation selection.**

(select one)

- Less than 2 years
- 2 – 5 years
- 6 – 10 years
- 11 – 15 years
- More than 15 years

• **I have worked on _____ digital preservation projects.**

(fill in the number of projects involved)

• **I gained most of my digital preservation selection experience in the following type of organization.**

(select all that apply)

- Archive (public)
- Archive (private) or corporate organization
- Library
- Museum
- Other memory institution (please specify)

• **I am a _____.**

(select one)

- Female
- Male
- Other

• **I am ___ years old**

(select one)

- Less than 25
- 25 – 44
- 45 – 64

[] More than 64

Part 2 – Selection Processes for Digitization

Topic A: What steps are taken in the digital preservation selection processes that archivists adopt in practice?

1. Can you describe the steps you or your team took in order to nominate and finally select a fonds or collection for a digital preservation project?
 1. Did you take these steps based on the recommendations of an existing guide?
 2. If yes, which guide and why was it chosen?
 3. If no, why did you decide on the steps that you did?
2. In the selection steps, did you discover that you needed to do some sort of decision-making or preparatory tasks that you did not anticipate or that were in addition to the commonly recommended steps?
 1. If so, why and can you describe these tasks?
 2. Did anything else surprise you about what needed to be done to conclude the selection process?
3. What kind of considerations were given to anticipating the way users of the digitized object would receive it?
 1. From an intellectual perspective (e.g. users will only look at portions of the content, or users will have pre-set expectations about the comprehensiveness of the digitized materials—thinking that it’s all the information that exists.)
 2. From a use-case perspective (e.g. historians frequently request this information, or a government task force was recently formed that will need information on this topic)
 3. From a hands-on perspective (e.g. access to the digital object could complement the original, or access to the digital object might replace accessing the original)

Topic B: How do archivists perceive the influence their personal roles (including experience, attitudes, and skills) have in creating the intellectual object that results from a digital preservation selection process?

4. How would you describe personal knowledge that you might have relied on to help you in the selection process? For example, a familiarity with common user needs that you developed over the course of your time working at your institution.
 1. Did your technical skills or your impression of the availability (or unavailability) of other experienced personnel within your digitization team influence your judgments during the selection process?
5. Considering your attitude or disposition on your own role and your institution’s role within our society, do you believe these roles should involve a greater or lesser engagement on issues in the public sphere?

1. Did you think that digitizing a particular fonds would influence events in the public sphere?
2. If so, do you think that influence was sympathetic with:
 1. Your institution's objectives
 2. Your own perspectives

6. Please think about the results of a project that you completed. With respect to the digital object that was created, does it match the expectations you had while selecting it?
 1. Is it being used as you expected, and if not, what is different from your expectations?
 2. Do you regret selecting it over other fonds that had been nominated? If so, why?
 3. Do you think that it is intellectually received (or understood) in the same way as its original?
 4. Based on your experience with the project and its results, do you think there are other ways to provide access to it (e.g. linking to other, potentially external sources of information, which complement the materials) that would affect how users perceive it?

Topic C: What are the recurring subjective factors in the identified processes that archivists make in their selections for digital preservation?

7. If you've worked on more than one digital preservation project, please describe how the selection phase differed between the projects, if it differed at all.
 1. Were any differences tied to your level of experience (e.g. a learning experience from one project may help with another)
 2. In what ways did you rely on personal experience to successfully repeat tasks in your projects?

8. Thinking about projects that addressed different user communities, did you find that your approaches in the selection phase were consistent? If so, what stands out in terms of your approach to the projects.

9. As an archivist (or whatever job function is relevant for the participant), do you believe that you have personal responsibilities that structure or influence the judgments you made?
 1. If so, what are these and how do they influence your judgments?
 2. Would you characterize these as well-known and understood to
 1. your peers and colleagues?
 2. your user community?
 3. the public sphere?

- 1 Association for Library Collections & Technical Services, Preservation and Reformatting Section, 2007, "Definitions of Digital Preservation". ALA Annual Conference Washington, D.C., June 24, 2007.
- 2 Michael Moss, "The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials Technical Guidelines for Digital Cultural Content Creation Programmes". *American Archivist*. 68, no. 1 (2005): 3.
- 3 Canadian Heritage Information Network, "A Guide for Managers Planning and Implementing Digitization Projects" rcip-chin.gc.ca, 2000.
- 4 United Kingdom, Parliamentary Archives, 2008, "Digitisation Guidelines".
- 5 Marc Fresko and Kenneth Tombs, "Digital preservation guidelines: the state of the art in libraries, museums and archives," European Commission, Directorate General XIII-E4 (1998), 41–47.
- 6 Moss, "The NINCH Guide," 40.
- 7 Maxine Sitts, *Handbook for Digital Projects: A Management Tool for Preservation and Access*, Andover, Mass: Northeast Document Conservation Center. (2000).
- 8 Canadian Council of Archives, "Digitization and Archives," (2002): 1.
- 9 Ibid.
- 10 Still Image Working Group, "Digitization Activities Project Planning and Management Outline. US National Archives and Records Administration," 2010: 11.
- 11 Ibid., 15.
- 12 Anna Bülow, Jess Ahmon, and Ross Spencer, *Preparing Collections for Digitization*. London: Facet Publishing, 2011, in association with the National Archives.: 49.
- 13 Ibid., 51.
- 14 Diane Vogt-O'Connor, "Selection of Materials for Scanning" in *Handbook for digital projects: a management tool for preservation and access*, edited by Sitts, Maxine K. (Massachusetts: Northeast Document Conservation Center, 2000), 46.
- 15 Vogt-O'Connor, "Selection of Materials," 48.
- 16 Moss, "The NINCH Guide," 40.
- 17 John McIlwaine, Jean-Marc Comment, Clemens de Wolf, Dale Peters, Borje Justrell, Marie-Therese Varlamoff, and Sjoerd Koopman, "Guidelines for Digitization Projects for Collections and Holdings in the Public Domain, Particularly Those Held by Libraries and Archives," (2002), 13.
- 18 Thomas James Connors, "The Notion of Value in Archival Theory: Some Lexical Archaeology" in *Who wants yesterday's papers?: essays on the research value of printed materials in the digital age*, edited by Yvonne Carignan. (Maryland: Scarecrow Press, 2005), 167.
- 19 Dan Hazen, Jeffrey Horrel, Jan Merrill-Oldham, "Selecting Research Collections for Digitization." *clir.org*. 1998. Council on Library and Information Resources., 83.
- 20 Hazen, Horrel, and Merrill-Oldham, "Selecting Research Collections," 83.
- 21 United Kingdom. The National Archives. Corporate Preservation Policy. [Richmond], 2009., 3.
- 22 UK National Archives, "Corporate Preservation Policy," 4.
- 23 Lorna Hughes, *Digitizing Collections: Strategic Issues for the Information Manager*. (London: Facet, 2004), 32.
- 24 Hughes, "Digitizing collections," 34.
- 25 Canadian Council of Archives, "Digitization and Archives": 1.
- 26 Joan Schwartz and Terry Cook, "Archives, Records, and Power: The Making of Modern Memory," 2002: 1.
- 27 Hilary Jenkinson, 1922, *A Manual of Archive Administration Including the Problems of War Archives and Archive Making*, 318.
- 28 Terry Cook, "Archival Science and Postmodernism: New Formulations for Old Concepts," *Archival Science*. 1 (1), 2001: 17.
- 29 Cook, "Archival Science and Postmodernism," 22.
- 30 Hazen, "Selecting Research Collections," 84.
- 31 Mark Roosa, "Some Thoughts on the Race against Time and Inherent Vice: Library Preservation in the Late Twentieth Century" in *Who wants yesterday's papers?: essays on the research value of printed materials in the digital age*, edited by Yvonne Carignan. (Maryland: Scarecrow Press, 2005), 29.
- 32 Roosa, "Some Thoughts," 31.
- 33 Janet Gertz, "Selection for Preservation in the Digital Age," 102.
- 34 Edward Hampshire and Valerie Johnson, 2009, "The Digital World and the Future of Historical Research 1." *Twentieth Century British History*. 20, no. 3: 396-414
- 35 Cokie Anderson, *Ethical Decision Making for Libraries*. Oxford, UK; Chandos Publishing Ltd, 2006.
- 36 Abby Smith, "Why Digitize?" *clir.org*, 1999, Council on Library and Information Resources.
- 37 Tom Nesmith, "Archivists and Public Affairs: Towards a New Archival Public Programming" in *Better off Forgetting?: Essays on Archives, Public Policy, and Collective Memory*, edited by Avery, Cheryl, and Mona Holmlund. (Toronto: University of Toronto Press, 2010), 123.
- 38 Allison Zhang and Don Gourley, *Creating digital collections: a practical guide*. Oxford: Chandos Pub., (2008), 22.

- 39 Robert Cole and Chris Hackett, "Search vs. Research: Full-Text Repositories, Granularity, and the Concept of 'Source' in the Digital Environment" in *Better off Forgetting?: Essays on Archives, Public Policy, and Collective Memory*, edited by Avery, Cheryl, and Mona Holmlund. (Toronto: University of Toronto Press, 2010), 123.
- 40 Joshua Sternfeld, "Archival Theory and Digital Historiography: Selection, Search, and Metadata As Archival Processes for Assessing Historical Contextualization,"(2011).
- 41 Greg Guest, Arwen Bunce, and Laura Johnson, "How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability". *Field Methods*, 2006, 18 (1): 78.
- 42 Maggie Jones and Neil Beagrie, *Preservation Management of Digital Materials: a Handbook*. London: The British Library for Resource, the Council for Museums, Archives and Libraries, 2001, 32.
- 43 Conway, "Preservation in the Age of Google," 69.
- 44 Cook, *Archival Science and Postmodernism*, 22.
- 45 Diane Vogt-O'Connor also identified a similarity between appraisal and selection for digitization. Vogt-O'Connor, "Selection of Materials," 45.
- 46 Verne Harris, "The Archival Sliver: Power, Memory, and Archives in South Africa." *Archival Science*. 2, no. 1-2, 2002, 84.
- 47 Stéphane Baillargeon, "La mémoire de l'art," *Le Devoir*, October 8, 2014, <http://www.ledevoir.com/societe/medias/423285/la-memoire-de-l-art>. (accessed October 9, 2014).
- 48 Hazen, "Selecting Research Collections," 83.
- 49 *Ibid.*, 84.
- 50 Nesmith, "Archivists and Public Affairs," 177.