Can We Do More?

An Examination of Potential Roles, Contributors, Incentives, and Frameworks to Sustain Large-Scale Digital Preservation

Carol A. Mandel

September 2019

Council on Library and Information Resources

Contents

Foreword: A New and Elusive Enclosure

Carol Mandel's cogent and sweeping introduction to her investigation of born digital preservation extends and enriches CLIR's long-held interest in protecting, making accessible, and sustaining our cultural heritage. A compelling story unfolds in this first chapter as we witness the design and implementation of frameworks to meet the demands of preserving the burgeoning human record, from the great library of Alexandria, focusing on Greek language and culture, to the florescence of national libraries, museums, historical societies, archives, and universities that followed. Each of these institutional types developed a well-defined conceptual focus (museums and physical artifacts, libraries and textual materials, archives and historical records) that, until the last three decades, almost exclusively pertained to physical formats.

These frameworks consisted of matrices of complex assumptions, rules, professional conduct, policies, and measures of success. One can imagine our cultural institutions as benevolent enclosures that reasonably satisfied the deeply human propensity to collect, preserve, and transmit the myriad instances of our creativity and invention. The physical objects of this stewardship also entailed certain properties and predictive characteristics. They were bounded in time and space in a way that could be described by affixed metadata that inferred continuity. But this is no longer the case.

Today we face a crisis that began several decades ago and is the impetus for Mandel's research: the appearance and explosion of born digital content, and the institutional budgets and policies, particularly in academic libraries, that have radically changed concepts of collection and preservation. Our current circumstances are deeply unsettling in the context of new digital content, and it is unclear how we will effectively manage the preservation of born-digital documentation—randomly iterative, unfixed, and ephemeral—with the professional tools and approaches we have inherited, or how we will redefine preservation as a core program within our institutions, or where we will find the needed funding.

"Can We Do More?," is an evocative question that engages our curiosity and insists on our looking to the future as we confront the truly wicked challenge of designing a new framework for a complex digital era. This first chapter provides a context for the work that must follow. The framing questions at the end of the chapter serve as guideposts to help articulate new methodologies. As the opening quote declares, while technology is a main source of our present tumult, the solution lies in successfully organizing ourselves to address this challenge. "Can We Do More?" Yes, unequivocally. But how we do it—the social, cultural, and behavioral realignments that this crisis requires—will define our legacy. CLIR is well situated and eager to engage wider communities, facilitate new alliances and collaborations, and help convene the expertise to resolve this crisis and establish a new, extensible professional framework for future generations.

—*Charles Henry*

"The past is never where you think you left it."
—Katherine Anne Porter (1890-1980)

"The problem of preserving digital information for the future is not only, or even primarily, a problem of fine tuning a narrow set of technical variables . . . Rather, it is a grander problem of organizing ourselves over time and as a society to maneuver effectively in a digital landscape."
—*Preserving Digital Information, Report of the Task Force on Archiving of Digital Information*, May 1996

Preservation and Digital Disruption

It has been 23 years since the Task Force on Archiving Digital Information ventured into unknown territory, boldly and brilliantly framing key challenges for the then-barely extant field of digital preservation. A generation later, the world of digital-only content has expanded in magnitude and modalities we could not then imagine. That born-digital world has given rise to extraordinary technical and professional development in digital preservation, thanks to an ever-growing number of talented participants who continually innovate to address new challenges. Many institutions, worldwide, are successfully capturing a rich range of born-digital content and managing large repositories of both digital and digitized material. A number of national libraries have stepped up to major initiatives, including the British Library's program of electronic legal deposit and distributed repositories, Australia's PANDORA web archive and digital collecting, and the acquisition of the Twitter archive by the Library of Congress (complete 2006–2017, and now selective). Diverse institutional and private initiatives have emerged to ensure enduring access to important content, from CLOCKSS and Portico for major scholarly material, to the ground-breaking Internet Archive for web sites and the Software Preservation Network, which seeks solutions to complex technical and contextual problems. The ever-growing number of organizations, coalitions, and well-attended professional conferences aimed at expanding expertise and action in the field of digital preservation demonstrates that the importance of this work is well understood—at least by a core of dedicated, knowledgeable professionals.

The last generation of development in preserving digital content has been transformative in the work of libraries, archives, and other cultural heritage institutions. But has it been transformative enough? The availability of digital-only content is certainly a "disruptive technology," both generally as it replaces familiar forms of documentary content and in the specifics of its endless variety of new modes and uses. That disruption is multi-dimensional, stemming from factors that go beyond the difficult issues of the technology, such as the design and management of repositories, processes for migration and emulation, addressing authenticity, and fixity. These factors include:

- the nature of networked content, such as social media and websites, which contain a continuing, often interactive, flow of posted material, discussions, and inventive combinations, disrupting known boundaries and forms of "documentary evidence"
- the vast scale of digital documentary evidence represented by everything from surveillance video to email to YouTube, requiring new perspectives on selection for long-term access
- the digital-only processes of creating, publishing, and distributing content in all forms, from research to entertainment to self-publishing, disrupting the methods of gaining a meta-view of the character and scale this work
- the intellectual property regimes that restrict the ability to collect and preserve so much substantive digital content, from copyrighted material to proprietary social media systems and news platforms, potentially requiring agreements with a complex array of private entities to acquire or rescue content for preservation
- the transient, ephemeral nature of digital-only content, eliminating the safety net and perspective of retrospective collecting, and requiring entirely new concepts and modes of predicting value and prospectively collecting

Organizations that collect, preserve, and give access to documentary content have been working to address these disruption factors for the last two decades, and there have been ground-breaking accomplishments. At the same time, the amount of digital-only content collected and preserved is small in relation to the scale of what exists. Few rational people would recommend collecting everything; while a case can be made for future interest in almost any particular form of evidence, e.g., as potential fuel for a "big data" analysis, feasibility demands some parameters for collecting. But have we collected and preserved enough? What is "enough"? Scholars continue to piece together knowledge of the earth and its civilization from random fossils and shards. Will future generations attempt to understand the early twenty-first century from the equivalent of digital ostraca, pockets of preserved content? As Clifford Lynch recently encouraged Canadian preservation collaborators, it is time to "pivot from focusing on the work of defining the mechanics of digital preservation to broad strategy" (CARL 2019). In the light of how deeply and thoroughly disruptive the world of born digital has become, "broad strategy" needs to extend to considering the mission and nature of the very organizations charged with the work of preservation, and to consider what is possible and what can realistically change.

The current structure of cultural heritage preservation institutions—libraries, archives, historical societies, museums, monuments preservation organizations—has evolved over several centuries, entirely shaped by a pre-digital world. It is a loose network of institutions with limited funding, local constituencies and expectations, and, while often collaborative, with largely uncoordinated strategies for collection and preservation. Through the end of the last century, coverage of the world's documentary heritage in its many pre-digital forms, while not ideal, had evolved to a predictable and acceptable pattern. This successful coverage was considerably enhanced by the ability to collect retrospectively: to take advantage of the efforts of an unusual, energetic private collector; to accept gifts of discoveries of papers in attics and storage sheds; to rescue materials abandoned by a former owner or a defunct organization;

and, most significant, to actively seek out materials previously considered unimportant and begin to collect them, such as personal papers illuminating African-American history or books such as cookbooks or graphic novels newly recognized as valuable primary sources. This is a pattern deeply challenged by a disruptive digital world in which so much content is truly ephemeral, not only because it exists in obsolete—but potentially excavatable—formats, but also because so much is locked into proprietary systems or is disappearing into discontinued platforms, lost accounts, discarded or erased storage media, and the general chaos of unmanaged and dispersed digital data. The collective resources of traditional memory institutions are probably insufficient to the full scale of the task, a scale made larger by a need to collect "just in case" because there will be no "just in time" for an unforeseen demand. In fact, "just in case" is more likely to be a become a need for "upstream" capture processes rather than traditional "downstream" gathering; such processes require new partnerships and new modes of working. At the same time, many memory institutions, particularly academic and public libraries, are focused less on collection and preservation and more on service and locally oriented programs, further stretching the resources available for new mandates such as digital preservation.

Society largely takes for granted that its heritage is being preserved by this reliable network of institutions. Cloonan cites a Gallup survey conducted in 1996 for the then National Institute for the Conservation of Cultural Property that found 95% adults polled agreed with the statement, "The collections in our nation's museums, libraries and historic houses need to be preserved" (Cloonan 2001, 234). That statement was evidence that the public values preservation; it can also be viewed as the expectation that those museums, libraries, and historic houses are serving as the nation's mechanism for preservation. While the act of collecting (or retaining in archives, or designating historic structures) is not synonymous with preservation, it is an essential prerequisite. Even as digital preservation professionals toil energetically and brilliantly to devise means to address technical preservation challenges, the outcomes of their work are necessarily limited by the capacities of their parent institutions to curate and collect born-digital content.

A generation ago, and for the first decade of the current century, the library profession imagined improvements in national policy and legislative changes that would facilitate and extend the work of digital preservation. Federal agencies have funded, and continue to fund, developments in technology and best practice. The Library of Congress National Digital Information Infrastructure and Preservation Program (2000–2016) contributed enormously to scoping and understanding how to address both the technical and organizational work of digital preservation. But it is clearly unrealistic in today's environment to expect legislative solutions, at least in the United States, or significant reallocation of limited resources by most cultural heritage institutions. In fact, to the extent that there has been additional funding and energy, it has focused on the digitization of important and endangered physical documents—a great boon to research, access, and preservation, but a potential diversion from collecting and preserving born-digital resources. The latter have not yet emerged as a pressing problem for most institutions. The challenge of preserving digital-only content needs, somehow, to be met within a feasible economic and institutional context.

In light of the disruptive challenges of the digital-only environment, it is time to re-examine the societal and institutional framework that currently serves to collect and preserve documentary evidence. At a minimum, such an examination can lead to a more broadly shared set of priorities and expectations about what should be captured and preserved to better focus future efforts; at best, it can identify additional players and resources to extend the capacity of society to preserve its story. This investigation begins by looking at the history and existing framework for the current preservation environment and considering where it does, and does not, match twenty-first-century challenges. The intent is to use this investigation and subsequent phases of work to frame issues for discussions by experts, leaders, and all interested members of the cultural heritage community. This chapter provides an initial framing that needs to be discussed; the aim is that these and subsequent conversations will lead to new approaches for considering particular kinds of content; outreach and engagement to expand the number and kinds of institutions playing a role in collecting born-digital resources; and the identification of targeted strategic initiatives where new avenues of activity could have substantial impact.

Preserving the Past: Knowledge, Evidence, and Cultural Heritage

Our current framework of institutions for preserving intellectual and cultural content stems from three distinct—albeit now interweaving and overlapping—traditions: the desire to amass and use knowledge, the protection of documentary evidence, and the embodiment of tradition and the past in cultural heritage. Broadly speaking, these traditions have formed the basis for today's libraries, archives, and museums, but that characterization oversimplifies the many kinds of institutions and organizations that collect, preserve, and give access to what David Levy calls "talking things," i.e., documentary evidence in its broadest sense (Levy 2016, 23). Understanding these traditions helps illuminate why so many kinds of hybrid digital "talking things" fall through the gaps between the missions of these different kinds of institutions.

Almost since its founding more than 2,300 years ago, the Library of Alexandria has symbolized the power of libraries to collect and preserve knowledge to be further used for the good of humankind. As Bagnall notes, "Within a century or so of its founding the Library had become a symbol of the universality of intellectual inquiry and the collection of written texts" (2002, 362). As described by Casson, "The policy was to collect everything [in Greek], from exalted poetry to humdrum cookbooks; the Ptolemies aimed to make the collection a comprehensive repository of Greek writings as well as a tool for research" (2001, 35). The Library of Alexandria can be seen as the source of the concepts of public library, research library, and national library, as well as the idea that a collection of texts can embrace, transmit, and preserve all of knowledge.

The actual "library" of Alexandria had a slightly different and broader mission than that of a modern library, as it was also famously the *Museon*, characterized by Casson as "a figurative temple for the muses, a place for cultivating the arts they symbolized . . . an ancient version of a think-tank"(33). The term *museon* is also the source of "museum," derived from collections of cuneiform tablets held in temple "libraries" (including temples to the muses). Writing about the intersections of libraries and museums, Martin reminds us that “In practice, there was little practical differentiation between a library and a museum until the early modern period, when the development of typographic printing resulted in a dramatic increase in the volume of texts available, which were then distinguished from a collection of objects, library from museum” (Martin 2019, 81). And while today museums focus on unique objects and libraries on books and documents, both have evolved with the mission of gathering and providing access to the “talking things” that transmit knowledge.

As evidenced in archaeological finds from earliest prehistoric times, today’s museums can attribute their origins to humankind’s ingrained propensity to collect objects of any and every sort. While many collections throughout history have been formed to accumulate perceived treasure, the museum collection is defined by the quest for knowledge, the combination of inquiry with acquisition, described by Samuel Johnson’s 1755 definition as a “repository of learned curiosities” (Simmons 2009, 1) and by the *Encyclopedia Britannica* as motivated by the goal of “preservation and interpretation of some material aspect of society’s cultural consciousness” (2019). As Simmons notes, museums are at their core “object-based institutions” (2009, 255). The *Britannica* begins its definition of museums by explaining that museums are distinguished from libraries by their collecting of *unique* items. As the modern museum emerged, institutions became more specialized and focused, using carefully selected, organized, and contextualized unique objects to bring knowledge to researchers and the public.

While museums were becoming more specialized, the intellectual scope of libraries was expanding. The growth of printed texts meant that libraries could accumulate books in large collections covering the widest imaginable span of intellectual content. This difference in scope between the collections of, often non-unique, content of textual and documentary material in libraries versus the focus on unique artifacts in museums has affected the nature and pace of digital disruption these memory institutions face. While digital technology has had substantial impact on the methods and reach of museums, the museum focus on using objects to impart knowledge has not—or perhaps more accurately, not yet—been reshaped by digital disruption. (Obvious exceptions to this are, or will be, the collection of digital artistic production or digital objects related to the history of technology and media.) The topic of potential expansion in the scope of collecting for museums in relation to the challenges of born-digital expression and representation will emerge later in this discussion.

Libraries and the Quest to Hold Knowledge

As printing and publishing became well established, libraries grew and flourished in the seventeenth century, and the concept of library collecting expanded to encompass the new scientific journals as well as books, the new form of texts. The fertile seeds of the great libraries established in the late eighteenth century in Europe and the nineteenth century elsewhere were planted and cultivated in princely and wealthy private collections of published literature. The modern realization of the Alexandrian goal to collect everything in Greek truly emerged in the formation of national libraries, tied to the concept of nation building in the Western world. Harris observes that a core conceptual view of libraries is that they “have always served to codify and preserve a desired sense of national identity" (2000, 6). With the British Museum, established in 1753, as forerunner, "central libraries with nationalizing impulses" arose across Europe and the Americas over several decades around the turn of the nineteenth century (Tucker and Goedeken 2009, 2085).

With the development of national libraries, the mission of knowledge preservation—within the bounds of national patrimony—becomes explicit. When UNESCO held two symposia to consider the roles of national libraries in 1958, it found enormous diversity in services and nature across countries. But at the same time, their report concluded that "in general, it may be said that the national library of a country is the one responsible for collecting and conserving the whole of that country's book production for the benefit of future generations" (Burston 1973, 183). In 1757, George II presented to the British Museum all of the books collected by the kings of England since Edward IV and also granted the museum the right to a free copy of all books published in the United Kingdom. When the Bibliothèque nationale was granted authorization to obtain manuscripts (1793) and then books (1795) from throughout France as a resource for the citizenry of post-revolution France, the notion that there were institutions charged with the public mission to preserve textual content as patrimony was born.

Through the nineteenth century, major public and university libraries were established with a goal to support intellectual development and research. In the United States, this impulse was exemplified by founding the Boston Public Library in 1854, the opening of the Newberry (1887), John Crerar (1894), and New York Public (1892) libraries, and the growth of collections at universities. As American universities began to model themselves on German research institutions, their libraries needed to build collections to support research. Joyce describes the late nineteenth-century period of development in university libraries as one that "brought a shift in values from conservation and protection of materials to access and use" (1988, 23) and which in turn created the need for separate areas to safeguard the rare and valuable items in those collections. Throughout the first half of the twentieth century, these segregated rare book collections became enriched by the acquisition of private collections, and then expanded into "special collections" that enabled libraries to broaden their capacity for collecting unpublished source materials, particularly literary and historical manuscripts that supported the scholarly "canon." Special collections expanded to include a variety of non-book printed formats with specialized subject focus in the 1930s (Joyce 1988, 24–26). The non-book materials were of wide variety—maps, broadsides, prints, sheet music—but of predictable subject scope for traditional historical and literary study.

At this point in time, a mere 70 years ago, the concept of accumulating and preserving the world's published knowledge seemed within reach, at least from a Western-centric perspective. Research resources that merited preservation for posterity were conceived as the published content that was being collected reliably by national libraries, and further enriched by the subject-oriented collecting of a few large and important research institutions. Non-book collections were growing, but were focused and manageable. And the acid time bomb that was about to embrittle almost all books created after 1850 was not yet recognized. This is an innocent moment and a useful point at which to take a look at the origins of protecting documentary evidence, i.e., archives and records, as the theories and traditions of knowledge and evidence will begin to interweave after 1960.

Archives, Evidence, and History

As long as humans have been writing, transacting business, and organizing themselves, there have been archives. The records kept in those archives were, of course, extremely important to the organizations that created them and to residents of jurisdictions that maintained evidence of contracts, deeds, and other agreements. In the West, the archives of the ancient world were largely destroyed; the maintenance of archival records began again in the eleventh and twelfth centuries to support emerging political and religious organizations (Duchein 1992, 15). The survival of archival records depended upon the level of care and attention given to them by their parent organization, and as such they were typically no less ephemeral than the organizations that created them.

Important as they were to so many aspects of life, archival records were not thought of as the basis for scholarship or knowledge creation until the sixteenth century, when the debate between Protestants and Catholics sent advocates to seek documentation of the origins and evolution of their respective churches—the same impetus that stimulated so much publishing during that period.[1] Even as archives became major public institutions—established as critical elements of the nation states that emerged after the French Revolution—they were viewed as quite distinct from libraries; their care and management was the responsibility of the organizations generating the records they held, even when some historical material was removed to a national library (Duchein 1992, 21).

While libraries sought to broadly embrace knowledge through collecting, traditionally archival collections were an outgrowth of the documentary production of their parent body. Historically, archivists charged with caring for organizational records were concerned with what to retain rather than what to collect. As the sheer scale of archival records grew, that framework required archivists to think deeply about the impact of their collecting and retention approaches not only on the present needs of their organizations, but on history. A foundational work on archival theory and practice, Hilary Jenkinson's *Manual of Archival Administration* (1937) arose largely out of the challenges posed by the masses of records that had been produced during the First World War (Tschan 2002). Jenkinson saw archivists as objective administrators who must base their decision-making about archival organization and retention on the values, priorities, and hierarchies of the parent organization, holding "the sanctity of evidence" above all else (Cook 2013). This point of view was soon challenged, however, by the growing use of archival contents for historical research.

The same intellectual trends and spirit of inquiry that gave rise to the development of book collecting and library growth during the late eighteenth century and throughout the nineteenth

[1] My thanks to Dr. Michael Stoller for providing erudite background on the introduction of archival records as authoritative evidence.

also engendered a strong interest in regional and local history and the active development of regional and local historical societies. Writing in the *American Historical Review* in 1934 (the same year that the U.S. National Archives and Records Administration came into being as an agency), Boyd describes the history of state and local historical societies in the United States from their origins in the *Zeitgeist* of the late eighteenth century (Boyd 1934). He notes that by 1884, the year the American Historical Association was founded, some 200 historical societies existed in the United States, fostering the work of enthusiastic local historians and collectors. The collecting interests of many of these societies were “heroic” (20), with the intent of gathering “materials calculated to serve a broad and inclusive definition of history,” or as described by Cox, operating on a “collect-everything” mission (2009, 209). While few could continue into the twentieth century with such broad ambitions, the research interests of their members and supporters demonstrated the value of a wide range of documentary types and topics—from newspapers and memoirs to topographical descriptions and meteorological observations—for understanding history. The enthusiastic and active work of these historical societies and similar organizations established the concept that an archive need not serve just as a reflection of a single institution’s operations, but also as a repository broadly conceived to provide “evidence,” whether originating as the arm of an administrative operation and serving its designated community or purposely shaped by collectors and specialists to serve the interests of their research.

That approach led another foundational thinker in archival theory, Theodore Schellenberg, to develop a new manual, *Modern Archives: Principles and Technique* (1957), emphasizing the view of the archivist as appraiser, an expert intermediary who used knowledge of historical practice to determine which materials are most important to retain for the future. With the introduction of this conception of archival retention, the theories of collecting employed in libraries and subject repositories were joined with those of archival retention to share an essential idea: the role of a professional mediator versed in historical perspective and methodologies who could ensure that the most needed primary sources would be preserved for future research.

Primary Sources

As Cook describes it, Schellenberg “faced the appraisal issue squarely in the mid-twentieth century in response to an avalanche of over documentation in all media” (101). But the concept of appraiser-to-the-rescue faced new challenges in the explosion of new and broadened fields of research that came with the enormous expansion of higher education and research activity in the 1960s in the United States and then globally. The interests of history became ever wider as social history embraced ever-more aspects of society and culture as topics of study. In response, libraries created new special collections and specialized repositories proliferated, exemplified by the Popular Culture History Center at Bowling Green State University; once popular culture became a subject of study, collecting demands became virtually all-inclusive. Content once considered outside the canon for serious study became the primary source documents essential for new literary and cultural studies, from dime novels to restaurant menus. As Cox describes, in a wide variety of repositories archivists began collecting “virtually every type of record representing every discipline, topic, and trend" (211). Beyond the records themselves, the meta-consideration of the nature and production of records—as evidence of business and trade, intellectual pursuits, leisure reading habits, and such—became an important element of study. The desire for an ever-wider range of primary sources has continued in academic research and teaching, with cross disciplinary approaches to inquiry and interpretation enabling limitless interests, and the importance of engaging students with the process of original discovery further increasing the desire for fertile primary sources (Hazen 2010, 3).

But as interests and documentation grew wider and more varied, collecting efforts could not be comprehensive, and instead, according to Cox, "contributed to the general randomness of collecting" (211). Even within organizational archives, the notion that expert professionals could predict what would be needed for future research became unrealistic. Cook summarizes the shift well in his description of how "the Schellenbergian model broke down" given the huge volumes of modern paper records and the transience of their new digital counterparts. "Moreover, historians were no longer the only serious users of archives; they were joined by scholars in many other academic disciplines, as well as lawyers, scientists, environmentalists, medical specialists, Indigenous peoples, human rights activists, teachers, artists, architects, novelists, heritage curators, lobbyists, and other professionals, and increasingly the citizen directly with expansive desires to connect with the past" (Cook 2011, 178–179). Those "expansive desires to connect with the past" converge with the third framework for considering needs in digital preservation: cultural heritage.

Cultural Heritage

The same movements, ideas, and historical circumstances that shaped the development of libraries and archives during the eighteenth and nineteenth centuries similarly drove the creation of museums for gathering and displaying physical objects, as well as stimulating the organization of efforts to preserve and restore architectural sites and monuments. Objects were collected from around the world to create a store of knowledge and, in the age of spreading colonialism, to demonstrate that these objects represented the wide reach of empire. Increasingly, the concept of national patrimony emerged and strengthened, exemplified in the nineteenth-century founding of the Commission des monuments historiques in France and its initiation of an inventory of buildings in 1837, a "canon" of important buildings. With the post-World War II formation of UNESCO and its subsequent World Heritage List, this canonical approach becomes globalized (West 2010, 33, 36). Smith has identified this official approach to determining valid and monumental cultural heritage as the "authorized heritage discourse," a terminology and bias issue now widely recognized in the cultural heritage community (2006, 4).

Three enormous changes in perception about cultural heritage have emerged over the last two decades, in part stimulated by wider global engagement in the activities of heritage identity and preservation. The first, identified by West as originating with the Burra Charter in Australia in 1999, was the introduction of social value as a criterion for determining the importance of a heritage site (i.e., not just the monumental attributes derived from aesthetics or historical events), including social value to a minority as well as a majority group. The importance of social value to understanding cultural heritage now extends well beyond physical sites and includes, among other things, documentary material. The legitimization of social value then leads to the second major innovation: the establishment of intangible cultural heritage, emerging in an "authorized" determination with UNESCO's Convention for the Safeguarding of Intangible Cultural Heritage passed at the 32nd session of its General Conference in October 2003. UNESCO's List of the Intangible Cultural Heritage of Humanity, also adopted officially in 2003, followed more than a decade of considering how to safeguard traditional culture and folklore. While this heritage may be "intangible," it is in fact captured and richly documented through sound, image, video, and even written descriptions. Once the limitations and bias of authorized heritage discourse were recognized, a third major concept, that of cultural self-definition, led to significant perceptual and operational change in how valuable cultural heritage is identified. As Smith describes, "Since the 1990s there has been increasing multidisciplinary interest in the way diverse communities forge, maintain and negotiate their identities. Alongside this, a range of communities, defined either geographically or by cultural, social, ethnic, economic and/or other experiences, have increasingly asserted the legitimacy of their collective identities and social, political and cultural experiences" (5). The recent understanding of the nature and importance of cultural heritage is inspiring; coupled with the capacity to be instantiated in myriad digital forms, it is a mind-bending challenge for collection and preservation.

In his thoughtful and much-cited essay on shifting paradigms in archival work, Cook considered the theoretical frameworks that have shaped the last 150 years of archival collection and retention. He characterized four phases of guiding thought: (1) evidence, the Jenkinsonian model of archivist guarding the juridical legacy; (2) memory, Schellenberg's widely adopted view of the historian-archivist selecting the archive; (3) identity, a post-modern view of a mediator-archivist shaping an archive to serve varied and broad societal perceptions and expectations; and (4) community, participatory archiving with the activist-archivist collaboratively mentoring community-based initiatives to create evidence and memory. The third paradigm, which developed in the 1970s during a period of rapidly widening academic and research interests, applies not only to archivists but also to collecting in libraries and other kinds of memory institutions that seek to meet society's vastly expanding expectations of what constituted sources for the pursuit of knowledge. The contemporary paradigm, community, is also now equally powerful for all memory institutions as they seek means to develop diverse and imaginative collections that no longer privilege an "official" perspective.

A second document passed at the UNESCO's 2003 General Conference was also potentially groundbreaking, although, as Karp notes, it "appears to have escaped similar notice" (2014, 159). This was the Charter on the Preservation of the Digital Heritage (2019), which emphasized the value and importance of born digital resources, declared that "the world's digital heritage is at risk of being lost to posterity," and called for urgent awareness-raising, advocacy, and the implementation of appropriate "legal, economic and technical measures to safeguard the heritage" by member states. This declaration makes its appeal to all memory institutions, and perhaps for the first time also points squarely to the role of museums in digital preservation. "The professional validity of museum concern with 'born-digital' materials would thus appear to be clearly established," affirms Karp, who also describes some of the challenges for museums in born-digital collecting (159). These challenges require significant rethinking and change, including the conceptual shift from a concern for preserving and accessing individual unique objects to managing digital objects that require replication, and the operational shift to entirely new infrastructures and processes for preservation (Karp 2014, 161). For a museum explicitly focused on community-produced intangible cultural heritage, the challenges are even greater, as some communities view social media as a key means for giving voice to their culture, from Facebook groups to expressive memes (Rees, n.d.).

Digital-only resources bring new challenges of collection scope for all memory institutions, and the challenges are not easily overcome. It may be telling that the thoughtful 2016 *UNESCO/PERSIST Guidelines for the Selection of Digital Heritage for Long-term Preservation* (Choy et al. 2016) aimed at helping archives, libraries, and museums implement the UNESCO digital heritage charter and that UNESCO's subsequent Memory of the World in the Digital Age declaration (Brungs 2013) have been cited only four times in Google Scholar and three in LISA—inclusive of several news announcements of their publication. In other words, it is safe to say there has not yet been a groundswell of adoption and activity.

Abundance and Constraint

As this very concise history of documentary collecting illustrates, the capacities of today's loose network of memory institutions are already challenged by the abundance of "talking things" and the diverse collecting objectives of the pre-digital world. While many specialized repositories and library special collections are healthy and active, their scope is necessarily bounded, even as they strive for more diverse coverage within their respective missions. They do not seek overall coverage, or even a larger coordinated pattern. They share information and expertise, and there are some successful collaborative—or at least non-competitive—endeavors in archival collecting, but as Cox notes, "such cooperation is a rarity" (215). The world benefits from the rich mosaic this varied and largely independent collecting activity creates, and there is advantage, even necessity, to such an environment, as it fosters preservation from multiple perspectives. The challenge may be less to achieve coordination but rather to encourage and ensure that there is *enough* collecting activity to cover a strong selection of the wide territory of valuable evidence.

The institutions that have historically been concerned with broad overall coverage are research libraries, the descendants of the Alexandrian ideal. But in the last few decades their collective coverage has become more circumscribed rather than extended as each library has necessarily responded to local priorities and constraints. In the 1980s, research libraries enthusiastically and optimistically sought to respond to the new demands of wide-ranging research interests through collaborative collection development. The Research Libraries Group (RLG) addressed this goal by developing the RLG Conspectus tool in 1980, a standardized method of assessing and describing collections. The Conspectus had many benefits and uses; the first two of these, as described by Gwinn and Mosher, were the need to identify collection strengths nationally and to enable "mutual reliance and interdependence in providing research materials" (1983, 130). The assumption was that libraries might give up collecting in areas where they were not strong, but would also maintain national-level, near-comprehensive collections in selected areas of strength. The Conspectus was widely used not only by RLG members, but also by library consortia across the United States and in other countries.

The tool proved extremely useful for understanding collections, but the dream of rich collaborative coverage was never realized, undermined in the 1990s by the converging local pressures of rising serials inflation and reduced library budgets. Spohrer describes the all-too-typical situation that existed at UC Berkeley in his article, "The End of an American (Library) Dream: The Rise and Decline of the Collection Development Policy Statement at Berkeley":

> It emerged as an outgrowth of the economic difficulties which the University faced in the early 1990s that the campus administration began to view the process of collaborative collection development chiefly as a means of "economization" . . . Cooperation for mutual enrichment of resources was transformed into cooperation as a means of relying on other institutions to meet a part of the basic collection needs of Berkeley scholars (2003, 41).

Assessing the use and benefits of the Conspectus, Atkinson reinforced that the tool itself had not failed, but was insufficient as a change agent:

> The reason we have been unable to achieve a broadly based cooperative program is quite simply that collection administration continues to be fundamentally influenced by that most subjective of all forces: politics. We have so far failed to counter those powerful internal pressures at all of our own institutions which demand that we devote all of our collection resources to the satisfaction of short-term, local needs. (1992, 354).

These pressures on and expectations for research libraries have only increased as they have transformed their roles to deliver an ever-wider array of support services for research, teaching, and learning in the last decade. While collection development librarians bemoaned these shifts in the 1990s, the exciting new roles for libraries and librarians in the digital campus are high priorities for universities. As UNC Libraries Director Elaine Westbrooks mused at a recent conference, "We're a research library, but we're shifting to service—that's where the students are" (ITHAKA 2018). Public libraries, both research and branch libraries, have undergone similar highly successful metamorphoses, with a wealth of well-loved community services taking prominence over collecting. The emphasis on service extends into the collecting mode, with much of the core collection made available through electronic access arrangements. This is a welcome and well-deserved golden age for libraries and their users, but it is not an age where acquisition and preservation define the library.

An understanding of these trends and realities is essential in considering feasible ways forward for born-digital collecting and preservation—an activity that is additive to all of the other competing expectations in today's libraries and repositories. But before turning attention to the digital environment, there is a success story that might also provide guidance in considering the future: the U.S. Newspaper Program. From 1982 to 2011, the National Endowment for the Humanities (NEH) led a cooperative national effort among the states and the federal government to locate, catalog, and preserve on microfilm newspapers published in the United States from the eighteenth century to the present. As described on the website (NEH 2019), with NEH funding and technical assistance from the Library of Congress, projects in all 50 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands were successfully completed. Each project was conducted by a single lead agency in each state, which in turn inventoried holdings in repositories throughout the state. The program had carefully crafted guidelines and procedures, clear lines of organization, sufficient funding, and universally shared enthusiasm for its goals. NEH and the Library of Congress are now leading a program to digitize a large segment of that newspaper content, aiming for about 300,000 digitized pages per state combined into a searchable database. These initiatives demonstrate that within a defined scope, a well-designed and well-supported preservation initiative with clear, shared goals can succeed.

The active National Digital Newspaper Program represents another important current thread of digital activity: digitization as a powerful tool for access to pre-digital content. From the seminal 1995 Making of America project to the dream of full-text Google Books through the European Commission's wide-ranging support for enabling Digital Cultural Heritage and Canada's Documentary Heritage Communities Project, initiatives of many kinds are exposing, compiling, and enhancing a vast array of specialized or historical collections across the world. The vision of public access to previously "hidden" collections has been compelling to funders as well as curators, and has attracted substantial support from government agencies and foundations of many kinds. These investments show that when a desirable, attainable goal is identified and a way to attain it is clear, collecting institutions and others in society can find resources to address it. An added impetus is the benefit of digitization as a scalable means for preservation and access, happily embraced as a replacement for microfilm for fragile paper materials and essential for rescuing rapidly deteriorating audio, video, and still image formats. The necessity of ensuring enduring access to these newly created digitized files has given rise to the development of preservation repositories in a variety of collecting institutions and to a diversity (some would say "complexity") of arrangements for digital preservation support services. While there are concerns that more institutions are creating digitized collections than know how to manage them to best standards (Fresa 2014), and while the determination of "best standards" is still evolving, emerging through this work is a baseline of investment and expertise essential to even contemplating the storage and management of born digital content.

Who are the Memory Institutions in a Digital-Only World?

The diverse players in today's complex of memory institutions are each considering—with widely varying levels of concern and commitment—the expectations of their collecting mission in relation to the vast and complicated universe of born-digital "talking things." Materials that were readily collected in their physical form now present challenges that require different expertise and processes, prompting reconsideration of collection scope. Maps have become geospatial data, newspapers are streaming content, photo albums are social media, and social media sites are the diaries of every person and every culture. Organizational archives and archivists are impressively stepping up to their obligation to manage their administrative and other organizational records for their designated communities; they are leading the way in addressing such difficult technical and policy problems as email and proprietary filing systems. Some national libraries are actively embracing their mission of national heritage preservation with programs of selective collecting from social media and web content and, ideally and to varying degrees, by extending legal deposit of publications in their country to e-only content. Research libraries have supported efforts such as Portico and CLOCKSS to preserve a core of the published research record and are initiating programs to manage local research data and faculty-generated unpublished sources. Library special collections and other specialized archives are learning to handle the digital-only formats of personal papers and small organizations, and some are beginning to capture relevant selections of web content. Energetic new archives are emerging to quickly document selected events and movements as they appear in broadcasting, social media, and websites. Archivists are enthusiastically exploring tools and practices to facilitate community archiving and personal digital archiving, and folklife centers and similar cultural heritage organizations now have digital ways to document intangible cultural heritage. And, moving much faster and well beyond traditional organizations, an extraordinary "private collector"[2] has created what many view as the single most important web collecting initiative to date: Internet Archive's Wayback Machine, launched in 2001. It is telling that the earliest, nimblest, boldest response to the challenge of born-digital collecting did not come from an established memory institution.

The question now is whether these many productive activities add up to a glass half full or one half empty. Does this potpourri of activity constitute a framework or foundation that will lead to a preservation enterprise that enables our descendants, or perhaps even ourselves when we are older, to learn about the world as it is today? There are still many areas that are worrisome. If we consider that the future will not allow very much, if any, retrospective collecting, will the current coverage of web archiving and selective social media mining add up at least to the kind of adequate historical record we now have in various special collections and archives for the pre-digital world? How will the largely ephemeral digital equivalent of personal papers be discovered and collected?[3] What record will we have of the e-only books sold on Amazon and not collected or even cataloged by libraries, or of online literary and poetry journals? What will happen to the proprietary music available only on the web? And perhaps the most disquieting question of all: we are drowning in current news coverage, but how much of it will be available to inform anyone who wants to interrogate it—and give it some much needed perspective—in the future? (Ringel and Woodall 2019).

[2] Clifford Lynch has aptly described Brewster Kahle as a passionate private collector.

[3] As so well described by Lynch, today's personal digital archives are vastly different from pre-digital personal papers (Lynch 2013).

This review is the first step in a series of explorations to frame issues and questions that can help shape shared strategies for future work. There is no one answer, and while there can be some leveraged investments—as are being made in digitizing collections and developing preservation tools—institutions will not be positioned to make major ongoing budget reallocations for activities that are not central to their mission. Essential elements of a broader collecting and preserving effort will need to continue to include a mosaic of material preserved by research libraries, especially national libraries, diverse special collections and subject archives, and community archives; these institutions need to be better positioned to do more. At the same time, we need ways to stimulate and facilitate local and regional collecting that are twenty-first-century versions of local historical societies and public library collections of hometown newspapers. We need to imagine how to bring the activity of personal digital archiving to entirely new levels and into new modes. We need to attract entrepreneurial engagement and innovative ideas to the challenge of digital preservation, to expand community-supported efforts such as Portico, CLOCKSS, and LOCKSS, and to create new ones. In a few key areas, such as streaming news and social media, we need to find workable centrally led (e.g., by a national library or a major consortium) approaches to the major challenges presented by critical proprietary content. And, as we identify the most promising initiatives, we may yet find ways to bring societal pressure to bear on large proprietary content holders and platforms—which are now feeling the pressure to be more socially responsible—to partner constructively with memory institutions. By framing issues, gathering answers to targeted questions, and structuring strategic conversations, we can take steps to leave a useful trove of digital knowledge, evidence, and cultural heritage for the near and distant future.

References

Atkinson, Ross. 1992. “In Defense of Relativism.” *Journal of Academic Librarianship* 17 (January): 353–54.

Bagnall, Roger S. 2002. “Alexandria: Library of Dreams.” *Proceedings of the American Philosophical Society* 146(4): 348–362.

Boyd, Julian P. 1934. “State and Local Historical Societies in the United States.” *The American Historical Review* 40(1): 10–37. Available at https://doi.org/10.2307/1838672.

Brungs, Julia. 2013. “UNESCO Vancouver Declaration: ‘The Memory of the World in the Digital Age,’” *Europeana Pro* (blog) March 25, 2013. Available at https://pro.europeana.eu/post/unesco-vancouver-declaration-the-memory-of-the-world-in-the-digi.

Burston, Godfrey. 1973. “National Libraries: An Analysis.” *International Library Review* 5(2): 183–94. Available at https://doi.org/10.1016/S0020-7837(73)80043-3.

CARL. 2019. “Digital Collections; Report of the Open Forum,” @Risk North: Collections in Peril, November 9, 2018, Montreal. Montreal: Canadian Association of Research Libraries. Available at http://www.carl-abrc.ca/wp-content/uploads/2019/03/AtRN2_report_EN.pdf.

Casson, Lionel. 2001. *Libraries in the Ancient World*. New Haven, Conn.: Yale University Press.

Choy, Sarah C.C., Nicholas Crofts, Robert Fisher, Ngian Lek Choh, Susanne Nickel, Clément Oury, and Katarzyna Ślaska. 2016. *The UNESCO/PERSIST Guidelines for the Selection of Digital Heritage for Long-Term Preservation*. UNESCO. Available at https://unesdoc.unesco.org/ark:/48223/pf0000244280.

Cloonan, Michèle Valerie. 2001. “W(H)ITHER Preservation?” *The Library Quarterly: Information, Community, Policy* 71(2): 231–42.

Cook, Terry Cook. 2013. “Evidence, Memory, Identity, and Community: Four Shifting Archival Paradigms,” *Archival Science* 13(2): 95–120. Available at https://doi.org/10.1007/s10502-012-9180-7.

Cook, Terry. 2011. “‘We Are What We Keep; We Keep What We Are’: Archival Appraisal Past, Present and Future,” *Journal of the Society of Archivists* 32(2): 173–89. Available at https://doi.org/10.1080/00379816.2011.619688.

Cox, Richard J. 2009. “Archivists and Collecting.” *Encyclopedia of Library and Information Sciences*, 3rd ed. Available at https://doi.org/10.1081/E-ELIS3-120044807.

Duchein, Michel. 1992. “The History of European Archives and the Development of the Archival Profession in Europe.” *The American Archivist* 55(1): 14–25. Available at https://doi.org/10.17723/aarc.55.1.k17n44g856577888.

Encyclopedia Britannica. 2019. “Museum | Definition, History, Types, & Operation.” Accessed August 29, 2019 at https://www.britannica.com/topic/museum-cultural-institution.

Fresa, Antonella. 2014. “Digital Cultural Heritage Roadmap for Preservation.” *International Journal of Humanities and Arts Computing: A Journal of Digital Humanities* 8: 107–23. Available at https://doi.org/10.3366/ijhac.2014.0102.

Gwinn, Nancy E., and Paul H. Mosher. 1983. “Coordinating Collection Development: The RLG Conspectus.” *College and Research Libraries* 44(2): 128–40.

Harris, Michael H. 2000. *History of Libraries in the Western World*. Lanham, Md.; London: Scarecrow.

Hazen, Dan C. 2010. "Rethinking Research Library Collections: A Policy Framework for Straitened Times, and Beyond." *Library Resources and Technical Services* 54(2). Available at https://dash.harvard.edu/handle/1/4111039.

ITHAKA. 2018. "We're a research library but we're shifting to services—that's where the students are." Tweet quoting Elaine L. Westbrooks at #ithakatnw18, Nov. 29, 2018, https://twitter.com/ITHAKA_org/status/1068259064688328704.

Jenkinson, Sir Hilary. 1937. *A Manual of Archive Administration*, new and rev. ed. London: Lund, Humphries.

Joyce, William L. 1988. "The Evolution of the Concept of Special Collections in American Research Libraries." *Rare Books and Manuscripts Librarianship* 3(1): 19–29.

Karp, Cary. 2014. "Digital Heritage in Digital Museums." *Museum International* 66(1–4): 157–62. Available at https://doi.org/10.1111/muse.12069.

Levy, David M. 2016. *Scrolling Forward: Making Sense of Documents in the Digital Age*, 2nd ed. New York: Arcade Publishing, 2016.

Lynch, Clifford. 2013. "The Future of Personal Digital Archiving: Defining the Research Agendas." In *Personal Archiving: Preserving Our Digital Heritage,* edited by Donald T. Hawkins. Information Today. Available at https://www.cni.org/wp-content/uploads/2013/09/Personal-Digital-Archiving-Cliff-Lynch-Oct-29-2013.pdf.

Martin, Robert S. 2019. "Intersecting Missions, Converging Practice." *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 8(1): 80–88. Available at https://doi.org/10.5860/rbm.8.1.281.

NEH. 2019. "U.S. Newspaper Program," National Endowment for the Humanities (NEH) website, accessed September 2, 2019. Available at https://www.neh.gov/us-newspaper-program.

Rees, Arran J. n.d. "What Does that Meme? Collecting and Curating Memes in Museums." Available at https://www.academia.edu/37932003/What_does_that_meme_Collecting_and_curating_memes_in_museums.

Ringel, Sharon, and Angela Woodall. 2019. "A Public Record at Risk: The Dire State of News Archiving in the Digital Age." *Columbia Journalism Review* (March 28). Available at https://www.cjr.org/tow_center_reports/the-dire-state-of-news-archiving-in-the-digital-age.php/.

Schellenberg, T. R. 1957. *Modern Archives: Principles and Technique*. Chicago: University of Chicago Press.

Simmons, John E. 2009. “History of Museums,” in *Encyclopedia of Library and Information Sciences*, 3rd ed., edited by Marcia J. Bates and Mary Niles Maack. Boca Raton, Fla.: CRC Press. Available at https://doi.org/10.1081/E-ELIS3-120044058.

Smith, Laurajane. 2006. *Uses of Heritage*. London; New York: Routledge, Taylor & Francis Group.

Spohrer, James H. 2003. “The End of an American (Library) Dream.” *The Acquisitions Librarian* 15(30): 33–47. Available at https://doi.org/10.1300/J101v15n30_04.

Tschan, Reto. 2002. “A Comparison of Jenkinson and Schellenberg on Appraisal.” *The American Archivist* 65(2): 176–95. Available at https://doi.org/10.17723/aarc.65.2.920w65g3217706l1.

Tucker, John Mark, and Edward A. Goedeken 2009. “History of Libraries,” in *Encyclopedia of Library and Information Sciences*, 3rd ed., edited by Marcia J. Bates and Mary Niles Maack. Available at https://doi.org/10.1081/E-ELIS3-120043544.

UNESCO. 2003. “UNESCO Charter on the Preservation of the Digital Heritage.” Paris: United Nations Educational, Scientific and Cultural Organization. Available at https://unesdoc.unesco.org/ark:/48223/pf0000229034?posInSet=1&queryId=f237fc01-5db8-446c-b477-6d675a19b34b.

Waters, Donald, and John Garrett. 1996. *Preserving Digital Information, Report of the Task Force on Archiving of Digital Information*. Washington, DC: Council on Library and Information Resources. Available at https://www.clir.org/pubs/reports/pub63/.

West, Susie. 2010. *Understanding Heritage in Practice*. Manchester: Manchester University Press.

Acknowledgments

This chapter, initially intended to be a brief introduction, took on a larger life as the need for deeper framing and analysis became increasingly apparent. I am indebted to many colleagues who reviewed and encouraged this work. My guiding lights have been Clifford Lynch, Jerome McDonough, and Donald Waters, who have been generous with their time and brilliant and insightful in their comments; their ideas have significantly shaped this investigation. Awesomely knowledgeable colleagues who have encouraged, advised, and reviewed my thinking at various stages of the process include Dale Flecker, Josh Greenberg, William Kilbride, Mary Molinaro, Richard Ovenden, Bernie Reilly, Oya Rieger, James Shulman, Michael Stoller, Winston Tabb, and Kate Wittenberg. I know they will continue to join the discussions and are working toward new strategies from many perspectives. I am very grateful to Chuck Henry for his inspiration and support; if this work finds an audience—and more important, some productive strategies and initiatives—it will be because CLIR is an organization well designed to stimulate innovation, problem solving, and engagement.

Framing Questions

Glass half full, i.e., what is being done/saved:

1. Are there successful current or recent collecting efforts that you think could/should serve as role models for others? What kinds of memory institutions could follow this model? What are the barriers (other than funding) that others would need to overcome to follow this/these models?
2. What do you see as key attributes of successful collecting efforts?
3. Which current or recent collecting efforts would be most beneficial if expanded? Or extended for similar but different areas?
4. Which current or recent collecting activities have the most potential to scale in a sustainable and affordable way?
5. How do we enhance the resilience and redundancy of current efforts across networks of memory institutions? Can we introduce redundancy in ways that reflect that some material is more important than others? How do we set these priorities?

Glass half empty, i.e., gaps and losses

1. What are we currently losing, or could easily lose, that will have the greatest negative impact on our future knowledge? What is or will be society's greatest memory loss? What concerns you the most and why?
2. What research questions, generally speaking, are being asked now that can't be answered because of lack of access to born digital content?
3. Who might be the new partners and players in caring about memory, knowledge, and heritage? Or at least in aiding memory institutions?
4. How should memory institutions change in the context of digital disruption? What should we expect from memory institutions?
5. Is the concept of memory institution still relevant?
6. When should institutions sample materials, when should they pursue curated collecting, and when should they strive to be comprehensive?
7. What strategies or initiatives can memory institutions undertake to address policy issues that affect collecting born digital content?

Organizational perspectives

1. If you work in a memory institution, or a related sector, does the collection (and thus preservation) of born digital content fit into your mission?
2. If there is born digital content that falls within your institution's collecting mission, what are/should be the highest priorities for your institution?
3. Have there been barriers to accomplishing those priorities? What are they? What might help your institution overcome them?
4. How do we measure success and failure?

5. Have you needed/will you need to reallocate human and/or financial resources to address born digital content priorities? How might that be managed?

Global perspectives

Digital disruption and its impact on future knowledge is a global issue. The author of this essay is writing from her North American perspective. What issues, concerns, or success models can you add from a different geopolitical perspective?