The impact of the shift to cloud computing on digital recordkeeping practices at the University of Michigan Bentley historical library

Dallas Pillen1 · Max Eckard2

Accepted: 24 April 2022 / Published online: 13 June 2022 © The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract
Cloud-based productivity, collaboration, and storage tools offer increased opportunities for collaboration and potential cost-savings over locally hosted solutions and have seen widespread adoption throughout industry, government, and academia over the last decade. While these tools benefit organizations, IT departments, and day-to-day-users, they present unique challenges for records managers and archivists. As a review of the relevant literature demonstrates, issues surrounding cloud computing are not limited to the technology—although the implementation and technological issues are numerous—but also include organization management, human behavior, regulation, and records management, making the process of archiving digital information in this day and age all the more difficult. This paper explores some of the consequences of this shift and its effect on digital recordkeeping at the Bentley Historical Library, whose mission is to “collect the materials for the University of Michigan.” After providing context for this problem by discussing relevant literature, two practicing archivists will explore the impact of the move toward cloud computing as well as various productivity software and collaboration tools in use at U-M throughout the various stages of a standard lifecycle model for managing records.

Keywords Records management · Cloud computing · Appraisal · University archives

Dallas Pillen
dallas.pillen@wayne.edu

1 Walter P. Reuther Library, Wayne State University, Detroit, MI, USA
2 Bentley Historical Library, University of Michigan, Ann Arbor, MI, USA
Introduction

Cloud computing has seen significant adoption throughout industry, government, and academia over the last decade. Cloud-based productivity, collaboration, and storage tools offer increased opportunities for collaboration and potential cost-savings over locally hosted solutions. The shift to cloud computing presents numerous challenges for records managers and archivists. As an example of this shift, the University of Michigan (U-M) 2016 Information Technology Strategic Plan states that “technology choices will favor solutions offered as external cloud services” in order to “accelerate the pace of research and innovation, cultivate interdisciplinary and inter-university collaboration, and drive economic development” (U-M 2016). This paper will explore the relatively recent history of the shift to cloud services in general and at U-M in particular, including its effect on digital recordkeeping practices at the U-M Bentley Historical Library (the Bentley), whose mission is to collect archival materials for U-M.

The shift to cloud computing

Cloud computing, which facilitates the use of "shared documents, files, software, knowledge, and applications... on demand" through the internet and remote servers, has seen significant adoption throughout industry, government, and academia over the last decade (Lin et al. 2014). Examples of cloud-based productivity, collaboration, and storage tools include Google’s G Suite, Microsoft 365, Box, Dropbox, and Canvas. The primary justifications for adopting cloud-based, rather than self-hosted, solutions are based around the increased collaboration afforded by cloud-based tools, which can be shared with and accessed by many users simultaneously, and also around the assumed cost-savings associated with moving from local administration to externally supported services. While these ostensible cloud-computing advantages benefit organizations, IT departments, and day-to-day users, the shift to remote infrastructure, administered by third-party entities and for which "most users do not have ownership," presents unique challenges for records managers and archivists (Pan et al. 2013, p 13).

In response to the growing adoption of cloud computing by record-making entities during the last decade, various record-keeping organizations and research groups have developed guidelines describing the risks associated with cloud-based record creation. A National Archives and Records Administration (NARA) bulletin, issued in recognition of the changes by which "records may be created, used, and stored" in the cloud, outlines several of the most pressing risks for archival institutions, including "maintaining records in a way that maintains their functionality and integrity," "maintaining links between the records and their metadata," and transferring archival records to those institutions or deleting temporary records at the end of their approved retention schedules (NARA 2010). A State Records of South Australia (SRSA) guideline further defines the
record-keeping challenges associated with cloud-computing, including "a loss of access to records," "record destruction or loss," and the risk that the "evidential value of records may be damaged" (SRSA 2015). The InterPARES Trust project produced checklists and reports outlining the potential issues that archivists and records managers should consider when contracting with cloud providers, including data ownership, availability, storage, preservation, retention, and security (InterPARES 2018). While some of these risks are not altogether unique to the shift to cloud computing, many are related to the fact that cloud-computing solutions tend to be administered by a remote organization whose infrastructure is not formally defined or made known to users, which increases the risk that data may not be maintained appropriately and that, when it is transferred out of the cloud service, relevant metadata may not be transferred along with the digital content.

The general shift to cloud computing throughout industry, government, and academia, along with its attendant record-keeping challenges, is exemplified by the adoption of cloud-based solutions at the University of Michigan (U-M). U-M is an R1, doctoral public research university in Ann Arbor, Michigan with very high research activity. In 2010, the university contracted a third-party company, Accenture, to conduct an IT assessment that led to a university-wide IT Rationalization project. The IT Rationalization project sought to address budgetary shortfalls following the Great Recession that affected information technology expenditures, stating that "business as usual is no longer viable for IT at Michigan" (U-M Office of the Chief Information Officer 2010). The IT assessment and rationalization recommendations led to the formation of the NextGen Michigan Collaboration Project, which had among its primary objectives "deliver[ing] IT in higher quality, less costly ways around campus" and "foster[ing] scholarly collaboration." In particular, the project sought to replace the "more than 40 e-mail and calendar services in use across campus" that led to "costly repetition and duplication of essentially the same services" (U-M NextGen 2011).

The NextGen Michigan Collaboration Project generated policies and reports that prioritized the adoption of cloud-based services, resulting in the implementation of several high-profile cloud-based services at the university. In 2011, the NextGen Cloud Strategy group recommended that the "University’s IT investment be focused on providing externally provided services that meet constituents’ high-level needs" (U-M NextGen 2012). Following these recommendations, the university adopted Google’s G Suite for Education collaborative tools in 2011 as part of "a growing trend toward cloud computing" with an explicit goal of adding "options for people to make their lives easier and allow them to work together more easily" (Goldberg 2011). In 2012, U-M implemented M + Box, an enterprise version of Box, to allow "individuals and teams to store, share, and discuss files" (Girardi 2012). By 2016, the university’s IT Strategic Plan stated that "Technology choices will favor solutions offered as external cloud services" among its IT Guiding Principles (U-M 2016). By the late 2010s, cloud-based collaborative tools were in widespread use throughout all levels of the university, with many records being created, used, shared, and stored in external tools such as Google Drive and Box.

Though archivists at the University of Michigan served on various NextGen project committees and drafted feedback documents in an attempt to take a stronger
role in the definition, identification, and preservation of records as the university moved toward cloud-based tools, many implementation details were made without direct involvement from archivists (McKay 2013). In hindsight, the use of such tools proliferated rapidly and the archival implications, though not fully realized at the time, have had a lasting impact on digital recordkeeping practices at the University of Michigan’s Bentley Historical Library.

About the Bentley historical library

The mission of the Bentley is, among other things, to “collect the materials for... the University of Michigan” (U-M Bentley 2020a) and for more than 80 years it has been home to the historical records of U-M. While the Bentley is also a major repository for archives and manuscripts relating to the history of the state of Michigan, the “University Archives preserves records that provide evidence of the policies and decisions chronicling the university’s growth and development, the creation of academic programs, research, student life, the physical campus, and the traditions that comprise the University’s 200-year legacy” (U-M Bentley 2020b). As defined in the U-M Standard Practice Guide (SPG), these “university records” are “all records, regardless of their form [emphasis added], prepared, owned, used, in the possession of, or retained by administrators, faculty acting in administrative capacities, and staff of university units in the performance of an official function” (U-M Standard Practice Guide 1997).

To paint the history of the University Archives and its approach to collecting the historical records of the University in rather broad strokes, it is helpful to point out that as an organization, U-M is, and prides itself on being, decentralized, and by state constitution independent in its recordkeeping practices. Even from its establishment in 1935, the Bentley recognized that not all was going to be centralized and proffered a “rather gentle definition” of what could be considered university archives (Bartlett 2001). In fact, in a survey letter dated just one year later in April 1936, the newly appointed Committee on University Archives had already come to terms with the fact that it was “manifestly impossible and undesirable to assemble all such records in one place,” and concluded that, “as one of its first acts, a useable catalogue of existing archives”—defined at the time as “all manuscript and printed material, official and unofficial, relating to the history of the University”—“and their location, should be compiled,” a revealing statement of a precustodial mentality (Ibid.). In a 2001 retrospective, Nancy Bartlett wrote that “today there is an irony for us in reading the language of the committee, since we can now realize that we have within the past sixty years moved from a precustodial phase of the university records here, through a largely custodial phase, to the beginnings of a postcustodial phase of a strategically planned, distributed records retention, due largely to the university’s digital environment,” as outlined above (Ibid.).

As it pertains to collecting university records, “regardless of their form,” as early as 1979 (Shallcross 2015), archivists in the Bentley’s University Archives and Records Program (UARP) were discussing the challenges posed by new technologies and machine-readable records, hosting a “Conference on Archival Management
of Machine-Readable Records” in February of that year (Geda et al. 1980). During what might be characterized as a foundational period in which the Bentley showed initial interest in digital recordkeeping, by 1997, the Bentley had received and developed a preservation strategy for its first significant collection of born-digital archives—the Macintosh personal computer of former University President James J. Duderstadt—and in 2000, initiated a web archiving program to capture snapshots of the websites of key academic and administrative units. The latter initiative, especially, was important in capturing the university’s response to the Y2K bug and the Grutter v. Bollinger Supreme Court case on the use of affirmative action in Law School admission decisions.

As the Bentley gained more experience and understanding of digital recordkeeping, it sought out innovative ways to provide access to these records and spearheaded such trends in the wider archival profession. In 2006, in partnership with U-M Libraries, the Bentley began providing online access to born-digital archives via DeepBlue, a digital preservation and access repository and U-M’s “permanent, safe, and accessible service for representing its rich intellectual community,” based on DSpace. In 2007, it hosted a two-day conference, “Development of Case Studies for the Effective Management of University Digital Records” (supported by an Andrew W. Mellon Foundation grant), which led to the development of the “Campus Case Studies” portal hosted by the Society of American Archivists.1

And, finally, as evidence that the new digital recordkeeping environment was ultimately transformative—or at least additive—for the Bentley, in 2009–2010, facing challenges of scalability and sustainability, a Mellon-funded “E-Mail Archiving at the University of Michigan” (or “MeMail”) project yielded a significant improvement in technical skills to curate born-digital collections in part because it included a dedicated “Technical Lead” position with sufficient time to research workflows and experiment with a wide variety of command line and graphical user interface (GUI) tools necessary to curate digital archives. This project also gave the Bentley first-hand experience with the ways in which the University’s highly decentralized “culture of choice,” while perhaps helping to attract and retain the “leaders and best”—as the University puts it—can also complicate records management in the new digital environment; the sheer number of email applications in use at the time across the University’s 19 schools and colleges (at the time, Outlook, Mac Mail, Pine, Mulberry, Web Mail, Thunderbird, etc.)—the same phenomenon that frustrated U-M’s IT during the NextGen Michigan Collaboration Project—made it difficult to find a single solution to capture important email (McKay 2015). In 2011, a new “Digital Curation” division was established with the goal of developing and implementing solutions for the long-term preservation and management of digital materials; later, it became the norm for processing archivists to process collections holistically, regardless of physical, analog, or digital format. This division

1 This portal, containing reports by university archivists who have created working solutions for a wide range of topics including managing born-digital records, collaborationing with institutional repositories, and developing records management policies for an institution, can be found at https://www2.archivists.org/publications/epubs/Campus-Case-Studies.
(and others) were ultimately brought into the fold of a larger “Curation” division whose approach to archival curation is holistic, i.e., using format-neutral intellectual arrangement and description. Archivists with curatorial roles were brought together in this new team from the Bentley’s previously separated divisions of University Archives and Records Program and Michigan Historical Collections. Finally, in 2014–2016 a Mellon-funded ArchivesSpace-Archivematica-DSpace Workflow Integration project united these three Open Source Software platforms in widespread use among North American archives to allow for the more efficient creation and reuse of metadata and to streamline the ingest of digital archives.2

All of which is to say that archivists at the Bentley feel fortunate to be among those archives with relatively well-established born-digital recordkeeping and digital preservation policies and procedures. These days, as the records of essential University functions—like conferring credentials, conveying knowledge, fostering socialization, conducting research, sustaining the institution, providing public service, and promoting culture (Samuels 1992)—are increasingly created and managed in various cloud-based productivity software and collaboration tools, they are adapting their digital recordkeeping practices accordingly.

Adapting the Bentley’s digital recordkeeping practices

While this paper’s primary focus is on records for which the Bentley has decided to take curatorial control over, the Bentley has also been involved in larger U-M conversations about records management writ large that provide context to the Bentley’s more recent digital recordkeeping history. As early as 2010, for example, the Bentley’s Director was engaged in campus-wide conversations about records management, including records management in the cloud. At the time, the director warned that cloud-based systems “entail a certain ceding of administrative control over the creation of records,” and that at least in the “near term, cloud space should not be used for highly consequential university records.” In 2011, they took part in a Records Management Task Force sponsored by the director of the Bentley, the Chief Information Officer, the Office of the General Counsel, and the Director of University Audits. This year-long project identified “issues of custody, control and access,” ownership (“Which office is responsible for maintaining the official copy?”), and the need for clarity around “self-provisioned”—i.e., not necessarily enterprise-wide—cloud storage. As it pertains to the latter, the Records Management Business Case the Task Force authored cautioned that the “widespread adoption of cloud-based services [was] driving a loss of control over university content.” Interestingly, the Task Force report also made far-reaching recommendations for the University, including a University-wide Records Manager, a new records management policy framework, and a new governance structure that would improve accountability and coordination of records management at U-M. The University did not ultimately adopt these

---

2 For more information on the grant project, see the initial press release at https://record.umich.edu/articles/bentley-receives-mellon-foundation-grant-develop-digital-archives-workflow/.

© Springer
recommendations, which is at least part of the reason behind the Bentley’s more narrowly scoped approach to records management detailed in this article. Internal documents from the time indicate that archivists recognized that, while there were some distinct advantages to cloud-based storage, there were also disadvantages, including “lack of control over contract terms with vendors,” “lack of trust in the future viability of vendors,” “lack of trust in the performance, security and availability of services,” and “the nonexistence of document and metadata management functionality.” It articulated a need for “exit strategies” as well as policy around what records could be entrusted to the cloud.

In 2016, the Bentley itself began to develop a formal records management program, envisioning a new “chapter” in the history of University Archives that built on its expertise at ingest and management of digital archival records and its strong relationship with campus units—strengthened all the more by programming around the University’s 2017 bicentennial—that included at least two roles: an Assistant Director for University Collections and Records Management with responsibilities for “planning and strategy for the records management program” and a new Assistant Records Manager, with responsibility for “training and development of retention and management guidance.” In particular, this records manager, hired in 2016, would assist units in the identification of “inactive records with long-term, archival value” and provide clarity and guidance for units using cloud storage including “how to create and store records, and how to destroy unneeded information or transfer records to the archives” (McKay 2016).

Since then, archivists at the Bentley have attempted to adapt to the shift to cloud computing on a number of fronts throughout the various stages of the model archivists there use for managing records: (1) creation (or receipt); (2) distribution; (3) (re)use; (4) management; and (5) disposition of those records with long-term, archival value (to the archives or to be destroyed). While a kind of gold standard, it can be difficult to have an opportunity to consult with units on the first stage of that model, the creation of their own records, except in an ad hoc way for individual projects. In recent years, however, the Bentley has begun to address the latter stages in more systematic ways, if not for all records created at the University, at least those for which the Bentley will eventually take curatorial control over. This has been an approach that, first and foremost, continues to take an active role in appraisal and selection, selecting records based on their long-term historical value regardless of where they’re stored (e.g., in local University storage, email, removable media, the cloud, etc.). In that respect, not a lot has changed simply because records these days are created or managed in cloud-based systems. However, this approach also:

- adopts some of what could be characterized as a more distributed, postcustodial approach to records management that attempts to provide guidance for units dis-

---

3 Of course, this model, documented at https://bentley.umich.edu/records-management/transfer-archives/, is only a model. It is worth noting here that there are other models, such as the Australian records continuum, that do not draw as sharp a distinction between where the creation and active management of records ends and the archival management of them begins.
tributing, using (or reusing), and managing their own records, including those on the cloud; and
• develops practical strategies for transferring select records from the cloud to the Bentley.

By 2019, in their “Records Policy and Procedures Manual,” they began to formally articulate the aforementioned guidance on cloud use from a records management perspective. They suggested, for example, that University units use the move toward a cloud-based collaboration service as a kind of “trigger” for the identification of inactive records and a determination of their final disposition, not unlike the advice they would give to units moving to a new building or experiencing a change in leadership. They add that “if you need to save space and decide to move towards a cloud-based collaboration service or an electronic recordkeeping system, we suggest spending the time you would have devoted to scanning on organizing and reviewing records to identify those [inactive] records that can be discarded or transferred to the University Archives.” Finally, they caution that, while cloud storage services can be great for real-time collaboration, they are not appropriate for inactive records with long-term, archival value the way that archives like the Bentley are, stating in their Records Policy and Procedures Manual that “services such as Google Drive, Box, and DropBox all offer convenient, user-friendly methods for storing data… but should not be the only location where critical records are stored,” adding that “care must be taken to ensure records are only accessed by those with the necessary permissions” (U-M Bentley 2019).

As a medium for how records are transferred to the Bentley, archivists there have likewise developed strategies for transferring data out of these platforms (e.g., using Google Takeout for G Suite), or using these platforms as a convenient means through which to transfer data stored locally. These are not without their drawbacks, like the fact that “cloud services, especially personal accounts, may or may not comply with laws, regulations or other policies for sensitive or legally protected information” (U-M Bentley 2020c). They’re also generally slower than other methods for larger transfers which characterize many of the Bentley’s recent accessions, whether relatively homogeneous runs of audiovisual archives (e.g., several years’ worth of raw video footage of campus events) or 1–2 TB transfers of 10 s of 1000 s of office documents (e.g., the personal “papers” of a politician and their office, these days what’s on their computer hard drives or office network drives). Some cloud services even have a limit as to how much can be uploaded or downloaded at a time. Moreover, “some significant properties (e.g., last modified times) of files may be overwritten when they are downloaded from the cloud” (Ibid.). Finally, in general, digital information in the cloud is not always captured in ways that are familiar to archivists. It is typically not moved into a centralized recordkeeping system, for example, and it is often difficult to attribute to one author and is not clearly linked to one business process or function. “Working with cloud transfers,” write Bentley archivists in their internal “Technical Transfer Guidelines,” can thus require “an additional layer of management,” especially since cloud storage does not “fill up” the way that a physical box or an external storage medium would (Ibid.). This is especially true for units that have established an ongoing relationship
with the Bentley and make annual or otherwise regular transfers; archivists must ensure that they are on the same page as donors about exactly what constitutes a given transfer.

None of these obstacles is insurmountable, of course. Archivists at the Bentley have developed strategies for addressing them—compiling multiple files into a single zip file before transferring via the cloud or any other medium can, for example, mitigate some issues with overwriting last modified dates and times—and have even adopted some of these cloud-based services into their internal workflows, at least for university records. One drawback of these university cloud-based solutions is that they are only available to university affiliates, even though the Bentley is also responsible for collecting records related to the history of the state of Michigan. Box (with some added steps to verify the integrity and authenticity of what was transferred), for example, is used in a transactional way for some of the most critical and sensitive records the Bentley collects, including the yearly transfer of positive and negative promotion and tenure casebooks (the latter of which is securely deleted after a given period of time according to a retention schedule) for teaching, research, and clinical faculty across all units, along with the unit’s promotion criteria. It is also used to provide individual researchers streaming access to digitized audiovisual material, some of which is copyrighted. That said, Bentley archivists (as well as their colleagues at U-M Libraries) have been intentional about not using cloud solutions when they are not appropriate, as is the case with the way they are approaching longer-term storage for records they have taken curatorial control over. For example, the Bentley’s Curation Team recently ensured that material they curate in long-term preservation storage no longer be moved to Amazon Web Services (AWS) cloud storage, and that content previously moved to AWS be restored to on-premises MiStorage provided by U-M ITS, precisely because cloud-based storage is not appropriate for this use case. While the ITS MiStorage service model includes moving (also called tiering) less frequently accessed files from a U-M managed data center to AWS, the purpose of Archival Storage is to “park” files there for the long term, in which case almost all content was or would have been moved to AWS. Since many of their preservation systems already use AWS as storage for backups (and only backups), this policy of tiering had the unfortunate effect of putting all of the Bentley’s “eggs,” as Lance Stuchell, Head of the Digital Preservation Unit at U-M Libraries put it bluntly, in one Amazon-branded “basket.”

**Issues with which archivists are grappling**

While archivists at the Bentley have taken steps in recent years to respond and adapt as best possible to this shift to cloud computing at U-M, the overall effectiveness of these steps has yet to be determined. They are currently grappling with any number of questions beyond the technical.

**The “snapshot” approach and significant properties**

Bentley archivists have opted for a “snapshot” approach—essentially the static capture of a dynamic resource at a given point in time—to transferring inactive
university records of enduring value from the cloud to the Bentley. This is not unlike archivists’ standard approach to “capturing,” on a set schedule, other dynamic types of records like websites and social media before compiling those snapshots into a more formal archive. When users make use of a tool like Google Takeout “to export… data to a downloadable archive file” (as described in Google’s documentation) what they actually get is a snapshot of some of that data as it existed on the day it was exported, and in a similar but usually not identical form.

Archivists are aware that for University units that use cloud computing services extensively—especially when they make robust use of those services’ feature sets—to manage their active records, this may not always be sufficient. Significant properties or characteristics of those records—while perhaps overloaded and ill-defined as a concept (Dappert and Farquhar 2009), those characteristics of digital objects that “must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record”—may be lost (Wilson 2008). So, for example, while Bentley archivists do try to ensure that the records that they download in this way are explicitly gathered or curated by donating units, and while it is possible for Bentley archivists to record facts about when the object was “captured” or “snapshotted” once they make their way into a preservation system that logs preservation actions, performs fixity checks, maintains provenance or a chain of custody, etc., it is still the case that some of the “story” behind these records, the collaborative aspects of applications like Google Docs that point to human relationships and interaction—like version history, suggestions, and comments—do not always make it into an exported PDF. Even if they were able to capture those aspects, the Bentley is ill-equipped to provide access to them in any meaningful way. Of course, this phenomenon is not unique to web archives or records management in the cloud. The embedded metadata in any given file, whether or not it was captured via the “snapshot” approach, does in fact become fixed at a certain point in time in a preservation system. However, it could be edited or altered—intentionally or unintentionally—before it comes into a preservation system, calling into question the trustworthiness of significant properties in the first place (as well as the evidential value of records, as in the following section) and centering the need for archivists and researchers to consider primary source literacy for all types of born-digital material. Admittedly, this would be even more critical if the Bentley were more concerned with records of artifactual (Society of American Archivists 2020), rather than evidential, value.

Evidence or process (or something else)

As Terry Cook argued in “Evidence, Memory, Identity, and Community: Four Shifting Archival Paradigms” (2013) there is a juridical legacy in archives, in which the archivist could be characterized as a passive curator and in which archival thinking is based primarily on archives as evidence. While this is shifting and transforming in the profession and at the Bentley (Cook also points out the alternative paradigms of memory, identity, and community), a key assumption of the “evidence” paradigm is that archives are made up finite, static entities—whether individual documents or
individual bits of data like “last modified dates” or “last modified times”—that can be pointed to, when needed, to support some assertion. This focus on evidence has arguably been an important part of recent archival educational opportunities around digital archives, for example, with its emphasis on digital forensics and the use of tools like Bulk Extractor that are borrowed from the law enforcement community. However, as Luciana Duranti points out, due to this shift to the cloud and the dynamic nature of tools optimized for real-time collaboration, the “by-products of such interaction are no longer finite entities, but processes that are always changing,” calling the evidential value of records into question (2015), even for those institutions like the Bentley with well-established born-digital recordkeeping and digital preservation policies and procedures. In some ways, these external cloud-based services are like “walled gardens,” or “non-open and corporate-controlled systems” that “prioritize[e] one type of openness (broad participation) over another (transparency)” by improving the ability of university-affiliates to use records in their active state while complicating the ability of archivists to faithfully preserve all aspects of those records in their inactive state (Squire 2017). These issues are not unique to cloud contexts and are also present when archiving locally controlled infrastructure such as organizational shared network drives (and, again, other dynamic types of records like websites and social media), but the opaqueness and lack of transparency around some cloud providers only serves to further complicate the issues. Indeed, while the “cloud,” as a term, can lead to a “somewhat misleading mental model which belies the physical reality of the vast server farms” owned by technology giants like Google, Microsoft, Amazon, and Apple, “from another perspective ‘cloud’ may be an apt term—as there are concerns about the extent to which the operations of service providers are ‘clouded’ in secrecy” (McKemmish 2013).

The “custodial” approach and the cloud

Cloud platforms—and associated ITS policies like the “tiering” example above—are optimized for active records and primary use, while archives are arguably optimized for inactive records and secondary use. Much of the above focused on the Bentley’s more-or-less custodial practices for collecting cloud-based records, especially the technology they employ for doing so. However, the issues surrounding cloud computing are not limited to the technology—although the technological issues are numerous—but also include organization management, human behavior, regulation, and records management; these factors make the process of archiving digital information in this day and age all the more difficult, especially in a highly distributed environment like the U-M.

As U-M has “outsourced” some of its storage technology and hardware for active records and primary use to various cloud platforms—and, given that there are access restrictions and contractual obligations from the cloud providers, it does seem that the material in these systems is on some level in the custody of the institution—the Bentley’s custodial approach to digital recordkeeping for cloud-based records has evolved. Archivists have adopted some practices of a more distributed, postcustodial approach to records management. This approach attempts to provide guidance for
units creating and managing their records, which is complicated by the aforementioned non-technological challenges, in order to ensure that those active, primary use records are able to eventually become inactive, secondary use records in the archive.

Policies aiming to control such behavior have not been successful and the University’s constitutional status also complicates the policy environment; while “subject to public records laws that mandate citizen access if demanded,” U-M is not “subject to the State of Michigan’s records retention schedule” (McKay 2016) so that there are no State mandates on public records or records management practices. In any case, academic archives like the Bentley do not have the authority to set or enforce such policies, and in some cases employers like U-M lack internal regulations on cloud computing use. At U-M, for example, ITS has provided ample documentation on what kinds of data are appropriate for what platforms, but little guidance on how best to manage cloud-based business processes or functions and the data that support them in the shorter- and longer-term. The Bentley has had the most success in providing guidance for units creating and managing their records specifically for records that will eventually be transferred to the archives.

None of this, of course, precludes an even more active role for the Bentley or other academic archives in the development and delivery of new records management and archival services for records in the cloud. “Archiving-as-a-Service” models (Stančić et al. 2013), for example, which call on archives and archivists to transition “beyond ‘custody’ from a ‘postcustody’ to a ‘postcustody 2.0’ paradigm” (McKemmish 2013) offer interesting paths forward for the archive to advise and influence records creators on their use of the cloud not only for active records but also for inactive records of enduring value. As McKennish notes, however, such models assume—questionably—that “best practice digital recordkeeping is already implemented,” that “good quality records and metadata are already being created and managed in digital systems,” and that “the organization’s recordkeeping and records are ‘cloud ready’” (Ibid.). At this point in time, it is therefore unlikely that the archivists at the Bentley would “advise” or “influence” records creators to place records identified as being of enduring value in the cloud.

Where things stand now

The cloud computing landscape at the University of Michigan continues to change in ways that further impact the recordkeeping practices at the Bentley. Recently, U-M announced that it will no longer be offering U-M Box as a storage service starting in December 2021 and will instead be moving to Dropbox, because the cost

---

4 In fact, a recent Michigan Supreme Court case has made this lack of State mandates on public records or records management practices all the more confusing for archivists at the Bentley, which in April 2021 deadlocked on whether records in possession of the University of Michigan library (i.e., the Bentley) are considered public record regardless of what guarantees the university made to a private donor regarding their release. See: https://www.courts.michigan.gov/4a4739/siteassets/case-documents/uploads/sct/public/orders/160012_73_01.pdf

@Springer
of using Box was going to increase significantly. In response to the announcement from a colleague, the Bentley’s current Archivist for University Archives reacted in this way:

[U-M] ITS has been pushing people to get rid of networked drives and move stuff onto Box. Lots of units believed them and changed their way of doing business. We will have to figure out a strategy for preventing a whole lot of deletion from happening in the next few months.

There is, perhaps, some irony in that part of the justification for both adopting Box and moving away from it was to save money. Regardless, such moves incur a kind of “technical debt,” a concept in software development describing the often hidden cost of additional work caused by choosing an easy solution now instead of a better solution that would take longer. There are real costs involved with units “chang[ing] their way of doing business,” as the Archivist for University Archives put it, costs that the Bentley, like all other units on campus including ITS, are now shouldering as their stakeholder representatives (including one of the Bentley’s stakeholder representatives, aforementioned Archivist for University Archives) work to gauge how Box is used at the Bentley to ensure as smooth a transition as possible away from it. To be fair, it appears that there is little risk that real data would be lost here, just ways of working and perhaps custom integrations with Box. In working with cloud technology providers, then, organizations like U-M are thus likely to regularly end up in situations where a different vendor providing a different infrastructure is likely to win a future contract which requires a major wide scale migration. It is worth noting that this is not fundamentally unlike what might happen with a wide scale refresh of underlying digital infrastructure even if an organization continues to manage its own data center; the migration of content between systems is, after all, a large part of the technology lifecycle for any form of information technology infrastructure. However, as Bentley administrators are quick to point out, units are left in fundamentally reactive positions, rather than in proactive ones that would enable them to better anticipate and perhaps manage these transitions, underscoring the essential need to create manifests and logs and provide support for organizational units to be able to transfer and migrate their material across those transitions from one cloud service provider to another. Otherwise it is likely that masses of working files may be lost that will ultimately result in the inability for that material to be transferred to the Bentley when they become inactive. This has the effect of disproportionately burdening units like the Bentley, who not only have a need—as do other units—to manage their own active records, but also have a mission to provide long-term preservation and access to university archives. The Bentley thus has to worry about the migration of their own data but also double down on efforts to provide the type of guidance to others described in the section above.

There is likely another paper to be written about the ways in which universities react to austerity-induced budget cuts by implementing cost-saving measures that, rather than acting as a panacea for cost and collaboration, further enrich large tech corporations and lead to increased maintenance or other deferred costs. Time will tell if the economic uncertainty caused by the 2020 COVID-19 pandemic results in similar rationalization efforts to those that followed the Great Recession. In the
meantime, as the university’s IT strategy continues to prioritize external, cloud-based record-making services, archivists at the Bentley continue to take steps to adapt to the shift in cloud computing for its recordkeeping activities. Though the overall effectiveness of these steps has yet to be determined, the Bentley is hopeful that its combination of practical approaches (for example, taking “snapshots” of records as they exist in cloud-based services using those systems’ available data-export capabilities) grounded in archival theory (for example, adopting a more distributed, postcustodial approach to guiding units on records management processes while maintaining an active role in the appraisal of records to be transferred to the Bentley) will ensure that cloud-created records that provide “evidence of the policies and decisions chronicling the university’s growth and development” will continue to be preserved.

Acknowledgements We are incredibly grateful for the contributions of our colleagues Nancy Bartlett and Aprille McKay that helped craft this paper.

Authors’ contributions Both authors contributed to writing the original draft as well as reviewing and editing that draft. Both authors read and approved the final manuscript.

Funding The authors did not receive support from any organization for the submitted work.

Data availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Code availability Software application or custom code sharing not applicable to this article as neither were generated during the current study.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Bartlett N (2001) Appraising university records: a retrospective analysis. http://www-personal.umich.edu/~deromedi/module/nbdisc.htm. Accessed 4 Dec. 2020
Cook T (2013) Evidence, memory, identity, and community: four shifting archival paradigms. Arch Sci. https://doi.org/10.1007/s10502-012-9180-7
Dappert A, Farquhar A (2009) Significance is in the eye of the stakeholder. In: Agosti M, Borbinha J, Kapidakis S, Papatheodorou C, Tsakonas G (eds) Research and advanced technology for digital libraries. Springer, Heidelberg
Duranti L (2015) Digital records and archives in the commercial cloud. In: Yoo C, Blanchette J (eds) Regulating the cloud: policy for computing infrastructure. The Massachusetts Institute of Technology Press, Cambridge, MA, pp 197–214
Geda C, Austin E, Blouin, Jr. F (1980) Archivists and machine-readable records: proceedings of the conference on archival management of machine-readable records, February 7–10, 1979. Ann Arbor, Michigan. Society of American Archivists, Chicago, IL
Goldberg H (2011) Gmail to replace University e-mail system next year. The Michigan Daily 26 October 2011. https://www.michigandaily.com/news/google-and-university-finalize-10-year-contract. Accessed 4 Dec. 2020
Girardi R (2012). M+Box accounts now available to faculty and staff at Ann Arbor campus. https://www.ur.umich.edu/update/archives/120409/mbox. Accessed 4 Dec. 2020

InterPARES Trust Project (2018). NA14: Trust in Cloud Service Contracts - Final Report. https://interparestrust.org/assets/public/dissemination/NA14_final_report_v5-1.pdf. Accessed 30 Mar. 2021.

Lin C, Yu WW, Wang J (2014) Cloud collaboration: cloud-based instruction for business writing class. World J of Education. https://doi.org/10.5430/wje.v4n6p99

McKay A (2013) Partnering with IT to Identify a Commercial Tool for Capturing Archival E-mail of University Executives at the University of Michigan. http://files.archivists.org/pubs/CampusCaseStudies/CASE-14-FINAL.pdf. Accessed 4 Dec. 2020

McKay A (2015) Capturing archival email of campus leaders at the University of Michigan. Archiving Email Symposium, 2–3 June 2015, Washington, DC. https://www.loc.gov/preservation/digital/meetings/aes15/5_McKay_20150602_LOC.pdf. Accessed 4 Dec. 2020

McKay A (2016) The future of records management at the University of Michigan. National Association of Government Archives and Records Administrators, 13–16 July 2016, Lansing, MI. https://www.sso.org/images/PDFs/NAGARA2016/C6_thurs_330-5_Where-Do-I-Belong_April.pdf. Accessed 4 Dec. 2020

McKemmish S (2013) Recordkeeping and archiving in the cloud: is there a silver lining? In: Gilliland A, McKemmish S, Stančić H, Seljan S, Lasić-Lazić J (eds) Proceedings of INFuture2013: Information Governance, 6–8 November 2013, Zagreb, Croatia. Department of Information and Communication Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia, pp 17-29

National Archives and Records Administration (2010). NARA Bulletin 2010–05. https://www.archives.gov/records-mgmt/bulletins/2010/2010-05.html. Accessed 4 Dec. 2020

Pan W, Rowe J, Barlaoura G (2013) Records in the Cloud (RiC) User Survey Report. http://www.recordsinthecloud.org/assets/documents/RiC_Oct232013_User_Survey_Report.pdf. Accessed 4 Dec. 2020

Samuels H (1992) Varsity letters: documenting modern colleges and universities. Scarecrow Press

Shallcross M (2015) A short(ish) history of digital curation at the Bentley Historical Library. http://archival-integration.blogspot.com/2015/04/a-short-history.html. Accessed 4 Dec. 2020

Society of American Archivists (2020) Artefactual value. https://dictionary.archivists.org/entry/artifactual-value.html. Accessed 4 Dec. 2020

Squire M (2017) Considering the Use of Walled Gardens for FLOSS Project Communication. In: Balaguer F., Di Cosmo R., Garrido A., Kon F., Robles G., Zacchiroli S. (eds) Open Source Systems: Towards Robust Practices. OSS 2017. IFIP Advances in Information and Communication Technology, vol 496. Springer, Cham. https://doi.org/10.1007/978-3-319-57735-7_1

Stančić H, Arian R, Milošević Ivor (2013). “Archiving-as-a-Service”: influence of cloud computing on the archival theory and practice. In: Duranti L, Shaffer E (eds) The Memory of the World in the Digital Age: Digitization and Preservation. UNESCO, pp 108–125

State Records of South Australia (2015) Cloud Computing and Records Management. https://archives.sa.gov.au/sites/default/files/documentstore/policies-guidelines/Standard/20150706_cloud_computing_and_records_management_final_v1.pdf. Accessed 4 Dec. 2020

University of Michigan. Bentley Historical Library (2019) Records policy and procedures manual. https://bentley.umich.edu/wp-content/uploads/2019/10/20191028_RecordsPolicyProceduresManual.pdf. Accessed 4 Dec. 2020

University of Michigan. Bentley Historical Library (2020a) About. https://bentley.umich.edu/about/. Accessed 4 Dec. 2020a

University of Michigan. Bentley Historical Library (2020b) Records management. https://bentley.umich.edu/records-management/. Accessed 4 Dec. 2020b

University of Michigan. Bentley Historical Library (2020c) Bentley Historical Library archival curation. Technical transfer guidelines. https://sites.google.com/a.umich.edu/bhl-archival-curation/accessions/technical-transfer-guidelines#TOC-Cloud-Transfers. Accessed 4 Dec. 2020c

University of Michigan. NextGen Michigan (2011). Project Update. https://wayback.archive-it.org/org-934/20110504204926/http://nextgen.umich.edu/projects/. Accessed 4 Dec. 2020

University of Michigan. NextGen Michigan (2012). NextGen Cloud Capabilities Team: Recommendations on Organizational Capabilities Related to Vendor-Based Cloud Services at U-M. https://wayback.archive-it.org/org-934/20140410042933/http://www.nextgen.umich.edu/cloud-computing/Cloud-Capabilities-Report-October-2012.pdf. Accessed 4 Dec. 2020
University of Michigan. Office of the Chief Information Officer (2010) IT Rationalization Fact Sheet. https://web.archive.org/20110514143357/http://nextgen.umich.edu/documents/IT_Rationalization_Fact_Sheet.pdf. Accessed 4 Dec 2020

University of Michigan. Office of the Vice President for Information Technology and Chief Information Officer (2016) University of Michigan information technology strategic plan. https://it.umich.edu/it-strategy/strategic-plan. Accessed 4 Dec 2020

University of Michigan. Standard Practice Guide (1997) Identification, maintenance, and preservation of digital records created by University of Michigan. https://spg.umich.edu/policy/601.08-1. Accessed 4 Dec. 2020

Wilson A (2008) Significant properties of digital objects. Joint Information Systems Committee Significant Properties Workshop, 7 April 2008, London, England. https://www.dpconline.org/docs/miscellaneous/events/142-presentation-wilson/file. Accessed 4 Dec. 2020

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Dallas Pillen is the Technical Services Archivist at Wayne State University’s Walter P. Reuther Library, Archives of Labor and Urban Affairs. He previously worked as the Archivist for Metadata and Digital Curation at the University of Michigan’s Bentley Historical Library and as an Adjunct Lecturer in Information at the University of Michigan’s School of Information. Dallas earned his MLIS with a Graduate Certificate in Archival Administration from Wayne State University’s School of Library and Information Science and is interested in archival systems, metadata, and digital preservation.

Max Eckard is the lead archivist for Digital Initiatives at the University of Michigan Bentley Historical Library. He has frequently published and presented on topics at the intersection of archives, digital libraries, and technology, including developing and teaching the Society of American Archivists’ Education course “Tool Integration: From Pre-SIP to DIP” and authoring the book Making Your Tools Work for You: Building and Maintaining an Integrated Technical Ecosystem for Digital Archives and Libraries. He earned his masters of library science with concentrations in academic and digital librarianship from the North Carolina Central University School of Library and Information Sciences.