Gordon Dodds Prize
Community Archives, Community Clouds: Enabling Digital Preservation for Small Archives

GRANT HURLEY

RÉSUMÉ Petits, locaux, indépendants et souvent singuliers, les nombreux centres d'archives communautaires de par le monde sont communément considérés comme les moins bien préparés à entreprendre des activités de préservation numérique. En me servant du concept des archives comme « endroit de préservation » et en le plaçant en dialogue avec des champs distincts de la littérature scientifique par rapport aux archives communautaires et à l'informatique en nuage, je soutiens que les solutions de l'informatique en nuage ont permis aux archives communautaires de commencer à acquérir et à préserver les documents créés en format numérique. Arriver à le faire nécessite de repenser les limites de l'endroit où s'effectue la préservation. Je présente les concepts qui définissent les archives communautaires dans le contexte anglo-américain, je discute de ces idées en relation avec préservation et endroit, et je les lie à un survol des solutions potentielles de l'informatique en nuage et à des études de cas pour des petits centres d'archives. Bien ancrés comme ils le sont parmi les créateurs de documents qui représentent une grande étendue d'organisation, de personnes et d'intérêts, les centres d'archives communautaires sont très bien placés pour acquérir les documents numériques de personnes, de sociétés et d'organisations qui pourraient ne pas être ciblées par les grands centres d'archives institutionnalisés. La nature distribuée des archives communautaires leur permettra de contribuer davantage à une approche de réseautage pour effectuer la préservation numérique, une approche qui reflétera la structure de réseautage de l'informatique en nuage elle-même.

ABSTRACT Small, local, independent, and oftentimes idiosyncratic, the many community archives across the globe are commonly considered the least prepared to undertake digital preservation activities. Taking up the concept of archives as a “place of preservation” and putting it in conversation with the distinct areas of scholarship on community archives and cloud computing, I argue that community archives are enabled by cloud-computing solutions to begin acquiring and preserving born-digital records. To do so means re-imagining the boundaries of where preservation takes place. I outline the concepts that define community archives in the Anglo-American context, discuss these ideas in relation to preservation and place, and connect them with a review of potential cloud-computing implementations and case examples for small archives. Embedded alongside records creators that represent a broad spectrum of organizations, people, and interests, community archives are ideally positioned to acquire the digital records of persons, businesses, and organizations that may escape
large, institutionally based archives. The very distributed nature of community archives will enable them to contribute to a more networked approach to digital preservation, one that reflects the networked structure of cloud computing itself.

Introduction

Archives of all types must preserve digital records if they are to continue to fulfill their social, ethical, and legal obligations to citizens and governments, and their broader responsibilities to memory, identity, and history. The question is no longer if archives will undertake digital preservation activities, but what records will be preserved, who will be responsible, and where preservation will take place. Decades of research have resulted in consensus on the key principles for digital preservation through theory and case studies treating the concept of the digital record, establishing authenticity, defining the needs of different records-creating organizations and individuals, and determining appraisal methods. The practice of digital preservation has also been extensively treated as agreement grows about the use and application of preservation standards, metadata schemas, and data models. The greatest unanswered questions are the interlocked issues of “who” and “where”: what scales of layered systems, software, hardware, and people should be maintained for digital preservation, and by which archival organizations? How much standardization is necessary so that these systems can speak to each other? As Ciaran Trace notes on the meaning of “born digital,” “archivists are trying to figure out what the appropriate nature and level of engagement should be with computer systems” prior to jumping in to the deep end of developing the necessary knowledge and infrastructure to actually do digital preservation as a routine activity. Some very large public archives have funded digital preservation systems at the national level, such as the US National Archives and Records Administration's Electronic Record Archive. The financial, technological, and labour resources required for these implementations are considerable, and the scope is as wide as the nation. Such approaches remain out of reach for many organizations, especially for the small, local, independent archives commonly referred to as “community archives.”

This article focuses on community archives to answer one aspect of the who and where of digital preservation. Taking up the concept of the “place of preservation” and putting it in conversation with the distinct areas of scholarship on community archives and cloud computing, I argue that despite the roadblocks of funding, technology, and training, community archives are enabled by cloud computing to begin acquiring and preserving born-digital

records. To do so means changing ideas about where preservation should happen while still protecting the values of the creating community and their trust in the safekeeping of community records. Archives as places of preservation are, and continue to be, defined by the bricks, mortar, and acid-free file folders of physical preservation, but cloud computing challenges archivists to reimagine the boundaries of where preservation takes place. Though a physical fonds can only be at one place at one time, the use of public cloud computing could spread the storage of records within a digital fonds across many servers around the globe. Archives will lose control over how and where their records are preserved if they do not claim and transform the space of the cloud as a place for preservation on their own terms. Extending preservation to the cloud offers community archives key advantages: the close connections that community archives maintain with their stakeholders, users, and donors means that they are ideally positioned to acquire the digital records of persons, businesses, and organizations that may escape the mandates and abilities of larger, institutionally based archives. The very distributed nature of community archives will enable them to contribute to a more networked approach to digital preservation, one that reflects the networked structure of cloud computing itself. This article focuses on community archives in Canada, the United States, the United Kingdom, and Australia, addressing the contexts of these institutions. I move from definitions of community archives to a discussion of the relationships between preservation, place, and cloud computing, and conclude by examining some potential cloud implementations that may work for small archival organizations. I treat archives at the organizational level rather than thinking about personal digital archives or web-based access and outreach projects, and I do not go into great detail about other aspects of digital records preservation, such as solutions for acquisition, transfer, and appraisal, as these subjects have been ably treated elsewhere.

While I intend the scope of this article to be as widely applicable as possible, I also write with a particular place in mind. My very first summer job was at the Charlotte County Archives, a small institution in St. Andrews, New Brunswick, that focuses on local history and genealogy. I spent four summers there processing fonds, digitizing photographs, and helping with research and reference. The archives are housed in a converted 19th-century prison adjacent to a neo-classical courthouse that was completed in 1840. From the courthouse’s stately portico and the jail’s forbidding granite construction, to the sloping lawns leading to the town’s commercial street and the dependably rising and falling tides on the shore below, the Archives seems to have a special kind of order and permanence that the chaotic metaphor of “cloud computing” at first appears to contradict. But the people and organizations reflected and represented in its holdings now create digital records as a matter of course, and cloud computing presents an opportunity for the Archives to acquire these born-digital records in step with its creators. Doing so means
creating a different kind of place for keeping records, one that is not as easily evoked as the Archives’ stone exterior. For this archives, and the many small archives like it, I bring cloud computing into the conversation with the hope that they can begin to act on digital preservation.

Definitions

The term community archives is used in this article as a matter of convenience to draw together what is in reality an incredibly diverse group of archival organizations. There are no strict features that can be used to define a community archives, and the word community has itself been treated with scrutiny by scholars such as Emma Waterton and Laurajane Smith, who argue that it can stand in for a wide array of social groups and constructions that are constantly in flux. “Community” does not have to be defined in terms of geography or ethnicity, though both of these sources of identity often motivate the creation of archives. Rather, community is a “frame of reference or orientation that coalesces around shared interests, common causes or collective experiences.”² Andrew Flinn notes that “community-led” or “community-based” are qualifying terms sometimes used to communicate that “control and agency” for an archives “is located within the group,” rather than being the initiatives of governments or other bodies that do not necessarily represent the interests of the community in question.³

The meaning of “community archives” varies based on the archival traditions of the relevant country. Community archives in the United Kingdom arose alongside the local government records offices that acquire the records of municipalities and other jurisdictions. Local history societies began acquiring private records from families and estates that complemented formal archival repositories, and many remain active in preserving these records today.⁴ Andrew Flinn and Mary Stevens have studied the many community archives established in the UK during the 1970s that seek to document evidence of marginalized groups, such as the histories of individuals of Asian or African descent living in London. Flinn defines “community archives” by the “grassroots activities of creating and collecting, processing and curating, preserving and making accessible collections relating to a particular

community or specified subject.”5 These organizations are frequently “motivated and prompted to act by the (real or perceived) failure of mainstream heritage organizations to collect, preserve, and make accessible collections and histories that properly reflect and accurately represent the stories of all of society”6 and sometimes position themselves as working against the mainstream archival institutions they feel have failed in this mission. In the United States, Australia, and Canada, community archives are generally considered closer in organization and mandate to the British local public authority archives, since they may collect municipal records or have more formal jurisdictional authority, but they also “share elements of the voluntarism and local focus of many UK ‘community archives.’”7 Such archives may be the creations of municipalities or public libraries, connected to First Nations organizations, religious orders, art galleries, or historical societies, or they may be entirely independent entities representing broad interest groups, such as the many archives acquiring materials from ethnic groups and LGBTQ+ communities. Different countries will have many variations on this theme as defined by unique histories and indigenous and immigrant populations. In Australia, Lyndon Ormond-Parker and Robyn Sloggett have discussed the community archives developed by many Aboriginal groups to reclaim their histories. These archives are “located not only in specific community archives but in the offices of the local Aboriginal land council, schools, language centers, art centers and more recently on computers within the community stored in purpose built ‘knowledge centers,’” developments that could include cloud-computing infrastructures.8 Stevens et al. suggest that community archives are worth defining by “concentrating less on the variety of terms in use and focusing instead on the type and purpose of the archives.”9 “Type” could include such factors as the goals and mandate of the archives; its financial resources; numbers of professional, para-professional, and volunteer staff; and what spaces, physical and online, are afforded for records under their care.10 At a minimum, a community must recognize its archives as such, and put efforts into consciously and

8 Lyndon Ormond-Parker and Robyn Sloggett, “Local Archives and community Collecting in the Digital Age,” Archival Science 12, no. 2 (June 2012): 196.
9 Stevens et al., “New Frameworks for Community Engagement,” 60.
conscientiously keeping records through practices that encourage responsible organization, access, and preservation, whether they acquire digital or physical records. Questions of purpose are equally interesting, as the different motivating factors for creating community archives expose different ideas of how community records should be acquired and maintained.

A second distinction is that of place and space. When I use the word place, I mean not only a physical and realized geographical location, but also the human actions and histories that exist in relation to that location, including economic, political, and social events. “Place” is both the coordinates and the corresponding experiences that are created, remembered, and sometimes recorded and preserved in connection to a location. As such, place is relentlessly particular to the landforms, people, and stories that constitute it. In contrast, space is a dimension in which matter is located: what that matter is, where it is, and why are not as important as observing the fact of its existence (or non-existence). It is defined by potential action and bounded by the rules or processes that constitute it. As a result, its study tends to flatten particularities to broader generalizations that focus on larger social or economic processes rather than specific incidents. Both of these concepts exist and act solely in relation to one another: there is no place without space and vice versa. The geographer Robert Sack explains the relation between place and space this way:

From the perspective of experience, place differs from space in terms of familiarity and time. A place requires human agency, is something that may take time to know…. As we move along the earth we pass from one place to another. But if we move quickly the places blur; we lose track of their qualities, and they may coalesce into the sense that we are moving through space.¹¹

Because the connections between people, events, experiences, and records are particular and situated, I see archives primarily as places rather than spaces. Because of its reliance on networks and lack of location, cloud computing is most often represented as a space. I see community archives as potentially reclaiming the digital spaces in the cloud as digital places. I turn now to discussing community archives as trusted places of preservation.

Community Archives and the Place of Preservation

In 2014, the Royal Society of Canada released an expert panel report titled The Future Now: Canada’s Libraries, Archives, and Public Memory. The report was the result of a year of public consultation and study to address the state of archives and libraries in Canada. One section of the report focused on the

place-centric purpose of community archives as informed by commentary from scholar Carol Couture on the “principle of territoriality, which asks that we make every effort possible to leave archives in the location where they originated.”12 The section concludes with the recommendation that “provincial and territorial ministries ... develop programs of financial aid that will allow communities to take charge of the preservation, treatment, and availability of archives.”13 No less important are the promises that identity or interest-based archives (including race, ethnicity, gender, sexuality, disability, or labour, among others) make to preserve the records of the individuals and organizations they represent. Sometimes their records may be acquired from larger areas than those covered by territorial-based community archives, but the principle is similar in that these archives act as independent places where particular communities entrust their records. The idea that community archives create strong reciprocal links between people, records, and place is a fundamental concept that writers in the past have linked to the value of these organizations. In 1953, Raymond Smith contributed an article on the British local record office to the manual *Local Records: Their Nature and Care*, in which he wrote:

The modern local archivist rarely limits his duties to the care of the official records of his Authority... The seeking out of records which ought to be in his charge, all undertaken in turn as his plan of campaign proceeds or as occasion serves, make him necessarily a mobile officer... Upon him lies a moral responsibility for providing against the destruction or dispersal of historical records which exist in private and public hands within his area.14

This statement applies today to many community archives: archivists working closely with individuals in their own communities can cultivate connections and relationships that lead to the responsible acquisition and preservation of community records in a trusted place. The ability of the archivist and the records in the repository to stand as evidence and memory for designated communities has been repeated by scholars such as Hugh Taylor, who wrote in a 1994 editorial that the value of small, community-based archives extends beyond “their great wealth of research material” to offering users “a sense of place in a constantly moving society,”15 by providing evidence and information on families, activities, persons, and businesses connected with that community.

13 Ibid., 88.
Similarly, Eric Ketelaar has suggested that records function as “memory texts through which the past is mediated”\textsuperscript{16} as a community encounters its archival heritage. Finally, Jeanette Bastian, in her 2003 study 	extit{Owning Memory}, argues that records creation and place or identity can be linked through the concept of a “community of records,” with the “community both as a record-creating entity and as a memory frame that contextualizes the records it creates.”\textsuperscript{17} Communities, geographically determined or otherwise, “create a mirror in which records and actions reflect one another in documenting the activities and forming the memory of the community,” which is then developed into bodies of records “preserved and valued by communities over time”\textsuperscript{18} as archives.

All of these concepts are bridged by the idea of archives as a “place of preservation,” which is advanced by Luciana Duranti in the 1996 essay “Archives as a Place.” Duranti argues for the centrality of the “archival threshold,” the “space where the officer of the public authority takes charge of the documents, identifies them by their provenance and class, associates them intellectually with those that belong in the same aggregation, and forwards them to the inside space”\textsuperscript{19} so that the records will stand as a “permanent monument to its creator’s action.”\textsuperscript{20} Though Duranti speaks generally about archival “space,” a realized threshold is contained within the singular place of every archives. The archival actions undertaken create recognition within the community of the archives as a trusted third party that can be counted on to preserve community records against alteration. Duranti’s perspective is at odds with some discourses on community archives that tend toward postcustodial approaches, which advocate that records should remain in the hands of their creators. Postcustodialism arose in the 1990s from writers such as David Bearman, who in his 1991 essay “The Indefensible Bastion” bids archivists to “give up their self-identification as the custodians of ... evidence.”\textsuperscript{21} The underlying assumption is that digital records could be kept more cheaply and with greater integrity by leaving them in the systems in which they were generated. Some scholars on community archives have taken up this approach in different ways. Andrew Flinn argues that “a postcustodial model is appropriate

\begin{itemize}
  \item \textsuperscript{16} Eric Ketelaar, “Sharing: Collected Memories in Communities of Records,” \textit{Archives & Manuscripts} 33, no. 1 (2005): 54.
  \item \textsuperscript{17} Jeannette A. Bastian, \textit{Owning Memory: How a Caribbean Community Lost Its Archives and Found Its History} (Westport, CT: Libraries Unlimited, 2003), 3–4.
  \item \textsuperscript{18} Ibid., 5–6.
  \item \textsuperscript{19} Luciana Duranti, “Archives as a Place,” \textit{Archives & Manuscripts} 24, no. 2 (1996): 244.
  \item \textsuperscript{20} Ibid., 247.
\end{itemize}
for community archives ... [because] it addresses the ambivalence that many communities feel towards depositing their archives in formal heritage institutions,” which implies that some community archives are themselves “postcustodial.”22 Terry Cook writes “that archivists “who work in the community ... [should] encourage archiving as a participatory process shared with many in society, rather than necessarily acquiring all the archival products in our established archives,” which takes up Bearman’s idea that certain records may not be acquired by archives at all.23 Examples of postcustodial community recordkeeping, such as the repatriation of records and artifacts taken from Indigenous communities, offer support for this approach.24 However, it is unclear whether postcustodial approaches to paper records in community settings also apply to digital records, since the primary principle of postcustodialism as advanced by Bearman is that digital records remain in their native recordkeeping systems, an idea that makes less sense for digital records created on personal computers outside of such systems. While archivists may collaborate with community members to advise on digital recordkeeping prior to acquisition, it is difficult to imagine multiple systems for long-term digital preservation being developed outside of a community archive’s jurisdiction, and it misses the point entirely to keep key records in the black holes of increasingly obsolete storage media or desktop computers. Furthermore, the archives and the creating community may be different parts of the same organization, with records simply transferred between the two, such as when a municipality transfers records to its town or county archives.

Recently, the discourse has turned back toward a custodial approach, which maintains that, pragmatically speaking, archives must support structures and resources for archival functions to occur if records are to have any meaning at all for the future. In her study of the Japanese American National Museum, Cristine Paschild writes provocatively that “the discourse of identity hinders the broader discussion of community archives in archival professional literature because of its inability to describe, address, and engage effectively with the realities and needs of the actual organizations themselves,”25 and that these institutions should above all ensure that archivists “not lose sight of our

commitment to preservation and access as fundamental services to all.”26 The visibility of archival functions, the designation of a place of preservation, and the position of the archivist who acts as a neutral mediator for communities create a set of “legitimating practices”27 that contribute to identity formation for the group that established the archives and also enable the community to care for its records. During his work organizing the Culion Museum and Archives, a community archives located on Culion Island in the Philippines, Ricardo Punzalan observed that “archival acts of arranging, boxing, labeling and exhibiting rendered the dormant records into meaningful archives that embodied heritage and identity.”28 The divides between these different discourses also rest on questions of space and place: the language of custody-oriented writers such as Paschild and Punzalan emphasizes the place-based relationship between communities and records over more conceptual or systems-based ideas about archival functions.

Archives as places of digital preservation is a concept that is just as important for community archives as it is for the archives of the state. Though the requirements for legal authenticity as evidence may be less crucial for personal or family papers, archival custody and proper processes for arrangement and description as well as preservation treat records with respect by ensuring they will not be lost or altered and can remain trustworthy memorials to the past. As many of the scholars cited have emphasized, a trust relationship between the community and its archives is fundamental to community identities and histories. The idea of an archives as a “place of preservation” is not conceptually limited to physical archives. In the second part of “Archives as a Place,” Duranti argues that it is equally important for electronic records to remain trustworthy in the future by making available in digital recordkeeping spaces structures that are similar to those we have in the built environment. However, the discussion of place and the assumptions of its relevance to community archives is challenged by cloud computing, both in terms of archives as places themselves and the geographical areas they can represent. Cloud computing is a technology that relies on obscuring place from view. As data stored in the cloud moves constantly between many “really boring-looking, windowless buildings with large, complicated HVAC systems,” in the words of technology writer Ingrid Burrington, connections to place are not made obvious to users.29

26 Ibid., 142.
28 Ibid., 208–9.
The next section investigates how archivists may use cloud computing in service to their communities to reclaim these digital spaces for themselves.

**Digital Records and Cloud Implementations**

The relationship between place and community archives, digital records and cloud computing complicates the issue of custody. Richard Cox comments in his afterword to *Community Archives: The Shaping of Memory* that “the current interest in community may be the result of so many thinking it is endangered in the digital era,” a sentiment that suggests “digital” is a spatial concept rather than a place-based one.\(^\text{30}\) It is no surprise that small, community-driven archives are experiencing a rocky period of transition to digital preservation. Conversations around small archives and digital preservation alternate from alarm to optimism. On one hand, recent studies warn that community archives are “overwhelmed”\(^\text{31}\) and “under-resourced”\(^\text{32}\); threatened with being left behind on the wrong side of the “digital divide”\(^\text{33}\) without adequate tools or knowledge; are unable to “invest in the technical infrastructure, resources and skills needed to become full participants in the Digital Age”\(^\text{34}\); and are subject to “lack of motivation”\(^\text{35}\) and “lack of attention”\(^\text{36}\) to the pressing needs of planning and implementing digital preservation programs. The sense of alarm is justified, since there are numerous roadblocks to digital preservation that community archives may be unprepared to face at once: the need for sustained knowledge and training regarding file formats, metadata, standards, and methods for digital preservation; the lack of stable financial resources to support planning, implementation, and maintenance; and the risks of committing to technical systems that require long-term support. On the other hand, initiatives such as the Preserving Digital Objects With Restricted Resources (POWRR) project, funded by the US Institute of Museum and Library Services\(^\text{37}\) and the National Digital Stewardship Alliance’s *Levels of...* 

---


\(^{32}\) Ibid., 28.


\(^{34}\) Ibid.


\(^{36}\) Ibid.

Digital Preservation,38 are intended to empower small institutions by providing the resources to begin preserving digital materials as “an incremental, ongoing, and ever-shifting set of actions, reactions, workflows, and policies.”39 Many resources are available to archives to provide guidance on digital records acquisition, transfer, and appraisal, such as Redwine et al.’s Born Digital: Guidance for Donors, Dealers, and Archival Repositories40 as well as the results of the InterPARES 3 research project, which was designed to offer practical approaches for existing bodies of records that are to be kept over the long term by archives with limited resources.41 New software systems such as Archivematica, Islandora, and Preservica offer varying levels of services that are beginning to lower the barrier of entry to digital preservation. I support the idea of “good enough” digital preservation as advanced by the POWRR project, which encourages institutions to try their hand at preserving digital records without spending vast resources. Accessible resources are becoming more common, including OCLC’s “Demystifying Born Digital Reports,” which gives readers easy-to-understand step-by-step instructions on solving basic digital preservation issues.42

Joan M. Schwartz and Terry Cook make the point that we may lose important digital records being generated in many contexts if “only certain types of information, and thus only certain people and organizations in society, are going to be privileged in our social memory by this new medium.”43 Creating opportunities for the preservation of born-digital records held by community archives will ideally provide a more accurate picture of our digitally mediated societies and the interests and experiences of individuals representing a wide spectrum of identities, ethnicities, and economic classes. Community

archives may also take advantage of their distributed nature to acquire born-digital records that might not otherwise be preserved. With hundreds of such institutions located in towns and cities in Canada and other countries, I see community archives acting as nodes in future networks of digital preservation services in a given country. This idea takes advantage of the fact that archives are already embedded around and alongside digital records creators: the task is now to develop usable and financially sustainable digital preservation solutions.

At first glance, cloud computing appears to be the antithesis of the place of preservation. Public clouds, the products most commonly available to users through services like Dropbox and Google Drive, depend on the placelessness of data. Digital records stored in public clouds, the most common cloud service available to consumers, can be located anywhere on any one of the servers within data centres owned by service companies, and data is moving among them constantly. In this framework, the cloud is space: potentially filled or empty and undifferentiated. It is the very movement of the data that enables cloud computing to be profitable, since storage can be constantly reallocated to take maximum advantage of the available space and computing power. As the authors of a 2009 paper on cloud computing note, it is quite ubiquitous (and is now even more so) but still appears to be a novel concept to users. In part, this is because cloud computing is based on the idea of networks and links between data within these networks, which is hard to conceptualize: “the cloud itself is an abstraction and is used to represent the Internet and all its complexity.” Since the word “cloud” implies something that is indistinct or opaque, users may also be unaware of the very physical aspect of data and data centres: somewhere, in some place, their data is physically manifested as electric signals on the hard drives of a server. The cloud does have a place, but that place is fleeting and obscured. A user’s records may be divided and backed up among a company’s data centres in any number of cities or towns across the globe, and among the numerous servers within these centres. Estimates of Amazon’s cloud services arm were done in 2014; it was speculated that the company owns 2 million servers spread across 11 global cloud regions and that most of their data centres contain between 50,000 and 80,000 servers.

The US National Institute of Standards and Technology (NIST) defines cloud computing as having five essential characteristics: on-demand self-service

(resources are available automatically at any time); broad network access (resources can be accessed using any connected devices); resource pooling (data dynamically moves between servers according to demand); rapid elasticity (services can be scaled or downscaled for any user to meet their needs); and measured service (use of services is metered and tracked for control and costs). The ability to spread data across many computers and rapidly provision services creates efficiencies in terms of server space and computing capacity, and benefits users by creating access to storage and computing abilities that would be costly and difficult to build otherwise. However, the rapid rise of cloud computing among businesses, governments, and universities means that jurisdictional and other legal issues relating to privacy, intellectual property, and research ethics are of paramount concern to users holding sensitive or private data. Questions about the search and seizure of servers by law enforcement, conflicts between the location of data and territory of storage sections in privacy legislation, and issues around security continue to concern users, governments, and policy-makers. All of these issues grapple with the problem of the cloud as both space (data can be anywhere) and place (where data is at a particular time).

Cloud computing is available to users in four different implementations: public, private, community, and hybrid. NIST defines public cloud infrastructure as “provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.” Public clouds are the most ubiquitous type, since they operate on economies of scale: the larger the network of data centres and critical mass of customers that use this network, the more value can be extracted from the built infrastructure that holds it all together. And though public clouds are the most commonly discussed, particularly in terms of the risks to privacy and copyright, their ubiquity belies the fact that other options are available through private, community, and hybrid clouds.

Private clouds are typically built by a single organization, though they can be shared with any number of users granted access in a model that resembles

---


traditional local access networks. The purpose is to offer some of the benefits of cloud services within an organization, such as making data and applications available to many different units, while maintaining full control over the storage and maintenance of this data. Some specialized third-party providers also offer private clouds where the user is granted exclusive access to a limited set of computing resources in a virtualized environment, or exclusive access to a physical set of servers.

Community clouds fall somewhere in between private and public clouds by sharing computing resources among a defined set group of users. NIST defines community clouds as infrastructures that are “provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations).” Access is restricted to this specified group of users, though implementations are typically larger than private clouds. Each member may contribute financial resources and servers to support the cloud, thereby increasing the efficiencies in the same manner as public clouds without sacrificing control. As in private clouds, community clouds can be set up by communities themselves, or they may choose to use services set up and operated by a third party.

Hybrid clouds are a fourth type of cloud implementation that may also see use. They are “a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability.” They typically take the form of a private cloud using the resources of a public cloud, or vice versa, and use information technology security techniques to create the storage “walls” between the two. Hybrid clouds allow organizations to use public cloud infrastructure for some data, and private or community clouds for other data in order to take advantage of the cost savings of public clouds. Where privacy considerations are crucial,
confidential data may be stored in the private section of the cloud, with an application connecting the confidential data to less sensitive data stored in the public cloud.  

The key difference between community and private clouds is the number of individuals involved in implementation and access and how many data centres are in use. Community clouds often have a series of distributed physical data centres, whereas private clouds make use of physical servers in a single place. In both cases, costs for these specialized implementations will be higher than for public clouds, and the outcomes of these decisions will depend on the needs and abilities of each archives or group of archives. Hybrid implementations may help lower costs by enabling archives to balance needs versus costs while still maintaining adequate control. For communities whose regions have the infrastructure, archives could make use of local data centres so that digital records are physically close at hand.

Based on the concepts relevant to community archives described in this article, the unmediated use of public cloud storage solutions is not an option, as cloud storage providers are unlikely to work directly with small community archives to protect their dearest digital records. Andrew Flinn describes the requirements for collaboration with community archives as “partnerships [that] are equitable, proceed from a position of mutual respect and recognition of the skills and expertise on both sides, and allow the organizations to retain their independence ... with knowledge and benefits flowing both ways,” a circumstance that public cloud storage is unlikely to offer. I believe that private and community clouds can enable community archives to build equitable and trustworthy relationships between archives, users, and the cloud, and fulfill the need to store digital materials in ways that are not as sharply divorced from the archives themselves. Community and private clouds can empower communities to maintain greater control over their records by creating clearer distinctions and greater transparency about where and how records are stored. Doing so claims cloud spaces as places that are particular to the archives that define them. Much like the structures in place for physical preservation, archives can set about defining rules and boundaries around the location of their digital records, determining what practices guide storage and preservation, and creating access policies. Since the establishment of many community archives is based on a set of shared values, community and private cloud implementations offer greater flexibility that acknowledges these values. For

56 Of course, built-in redundancy is necessary to back up data at other data centres that account for natural disasters or other risks.
example, the *Aboriginal and Torres Strait Islander Library and Information Resources Network Protocols* recognize that some Indigenous cultural knowledge is secret and sacred and should not be made generally available to the public, or is to be accessed only by selected individuals within their own community.\(^{58}\) Such materials may take digital form and require appropriate handling. Scholars Jane Anderson and Kimberly Christen have developed Indigenous knowledge licences and platforms that respect these protocols and requirements for “safe keeping places” of important knowledge.\(^{59}\) If preserving these types of materials in a cloud is to be viable, implementations must follow protocols to create “trust that is built around respecting the ethical and normative systems that already exist within Indigenous communities.”\(^{60}\) Cloud implementations also enable an array of new interactions with archival records. Communities can control access to digital records at any time through multiple interfaces. Records could be linked through the community archives’ website, local and regional consortia search portals, and national or international catalogues. Opportunities for tagging, sharing, and commenting on records may also match functionality in social media applications that are now commonplace, and new functions specific to archives such as collaborative arrangement and description, flexible aggregations, and personal collections of records developed by users could be integrated into these interfaces. These alternative implementations of cloud computing make room for community archives in these often intimidating digital spaces, and allow archives to extend the boundaries of their institutions to digital places they can trust.

**Possible Implementations: COPPUL and ArchivesDirect**

The greatest challenge to working with community and private clouds is that greater control over the system means greater costs and complications: Amazon’s and Google’s cloud services are profitable for that very reason. One attractive possibility for community archives in the cloud is the creation of consortia clouds where smaller community archives pool resources to store their records in a centralized private cloud. Such an approach would allow archives to demonstrate that records are being kept in a trusted place while reducing costs and technical barriers to digital preservation. Communities holding physical records have already taken advantage of similar models for custody; for example, communities may choose to transfer records to archives with better preservation capabilities, using permanent or negotiated

---

59 Anderson and Christen, “‘Chuck a Copyright on It,’” 111.
60 Ibid., 112.
loan periods. The approach to digital preservation adopted by the Council of Pacific and Prairie University Libraries (COPPUL) in Canada offers a strong example of how consortia of interested institutions might collaboratively develop a cloud solution. Hosted on the University of British Columbia's EduCloud service, the Archivematica as a Service approach allows for member universities to use a cloud-hosted instance of the digital preservation software Archivematica without having to install the software and set up the necessary storage infrastructure on their own. The servers used reside at UBC in a private cloud. A second Canadian cloud provider called OVH is used as a backup and testing environment. Therefore, organizations can access digital preservation services without the hurdles of infrastructure and technological expertise, balanced by membership costs and the need to ensure that local staff has an adequate knowledge of digital preservation. Project members Bronwen Sprout and Mark Jordan note that one of the difficulties experienced in implementing the projects was not a lack of technical knowledge but a lack of “comprehensive digital preservation policies or frameworks,” which the project motivated participating institutions to develop. As the panel that authored the recent Council of Canadian Academies report Leading in the Digital World: Opportunities for Canada's Memory Institutions notes in relation to the COPPUL project, the ideal outcome is that more consortia-based projects for digital preservation will lead to larger groups of nationally aggregated systems that will require greater agreement on digital preservation standards, including choice of metadata and file formats.

62 Archivematica is a set of web-based micro-services that enables digital preservation through defined processes built into the system. It is based on the Open Archival Information Systems model. Preservation workflows ingest new digital records, normalize files to preservation formats, and create preservation and access packages with PREMIS and METS metadata. Over time, files can be migrated to new formats to combat format obsolescence while maintaining the integrity of the originals. The system is open source, and preserved files are independent from the system itself, meaning that users can migrate their digital content elsewhere. For a good overview of the system, see Brad Houston, review of Archivematica, American Archivist Reviews, 2 March 2015, http://www2.archivists.org/sites/all/files/Archivematica.pdf.
64 Ibid., 242.
A second approach, called ArchivesDirect, was released in March 2015 through a collaboration between Archivematica and DuraSpace. DuraSpace is a non-profit organization that has been developing institutional repository systems such as DSpace since 2002. The approach integrates Archivematica with the DuraSpace’s DuraCloud application. The partnership means that users can deploy Archivematica without the need for local server infrastructure, much like the COPPUL approach. Users access Archivematica through a web dashboard. The difference between the COPPUL approach and ArchivesDirect is that the latter works using a public cloud implementation, with DuraCloud functioning as an intermediary between the user and the cloud. DuraCloud is an application that works on top of Amazon’s S3 and glacier public clouds in order to overlay preservation and access capabilities. The system stores content on the two clouds at once, to mitigate the risk of one provider losing availability or files. It also synchronizes the files between providers, performs health checkups for integrity using checksums, and allows users to move files to and from providers. By combining Archivematica with DuraCloud, users are able to take advantage of the digital preservation capabilities of both systems. However, the disadvantage of ArchivesDirect is that because the service uses public cloud infrastructure, it may not be possible for archives working in jurisdictions with territory of storage privacy legislation to use it for all of their data, and a hybrid solution may be necessary.

Archives considering both types of implementations will have to evaluate what kinds of records need to be preserved; where these records can be preserved according to the practices, protocols, legislation, and values supported by their communities; and what other kinds of financial resources, technical infrastructure, and knowledge might be needed to support either system. Despite the promise of consortium-based cloud services, the barrier to digital preservation will not immediately disappear for small archives. A basic level of theoretical and technical knowledge in digital preservation is required to operate Archivematica. Knowledge of the methods required to ensure proper transfer and ingest of digital materials in different file formats from donors must be also considered. Since digital records will not usually be captured from existing records systems at the time of creation, strong descriptive metadata relating to provenance and the technological context of original digital records is ideal. Finally, financial resources will have to be invested in locally operated computers to connect to the cloud. Membership fees in a consortium may help fund solutions to these issues, but the cost of these may remain too high for the smallest of archives. A sliding scale based on institutional

---

resources is one approach to ensure accessibility. Similarly, the existing infrastructure of provincial archives councils and associations may one day be able to provide training and support to cloud newcomers. There are no simple solutions to these hurdles, but the many accessible resources described above, and some new creative and collaborative approaches, will help lower them.

**Conclusion**

Returning to the concept of the place of preservation, community archives (and the archival profession generally) must consider how they define place in relation to their digital records. If an archives is trusted by the community as a place where preservation occurs, then these questions remain: How is place conceptualized? Where should a digital archival space be? In the past, physical records had to be preserved in a particular building with a realized geographical location. Place could be easily understood in relation to archival records: a fonds or collection took up physical space on a shelf, and the risks of destruction or decay were mitigated by the archivist using established standards for physical preservation. The use of cloud computing for preservation challenges this concept because of the separation of the digital record from the place where it is stored. The case of interest or identity-based community archives without geographical affiliation (such as a transgender-focused archives that acquires records regardless of their place of origin) offers some insight into this problem, since the records held do not necessarily relate to the place in which they are stored. In these instances, the archives could theoretically be located anywhere: the crucial piece is that the supporting community understands and trusts the archives as the unique place of preservation and that the stories these archives tell are unique to that community. This particularity is what makes an identity-based archives as much a place as a territorial-based one. Community archives of all types can preserve digital records in the cloud with the same logic. Digital spaces must be reclaimed by community archives as digital places where interlocked standards, practices, policies, and technologies unique to the needs of the community protect their records. In so doing, communities can actualize the place-ness of digital networks that are usually obscured from view. These structures are no less built than physical archives, though they are harder to imagine than the romantic picture of dusty archives. But imagine we must, so that in the eyes of their communities, archival places for preservation, whether digital or physical, are seen as equivalent and interchangeable. Just as they have in the past with physical archives, communities can use the collaborative nature of cloud computing to claim their own right to its use and to assert the real importance of preserving their records.

Some archives may decide that private or community cloud implementations create the required links between the archives and digital place and justify the added cost, whereas others may be comfortable using public
clouds if the community itself feels that an adequate measure of trust can be produced by this relationship. I support the idea that public and community cloud implementations, or mediated public cloud services such as that available via ArchivesDirect, can create trusted places of preservation for community archives. Since the expensive and complicated infrastructures for formally certified trusted digital repositories are far out of the reach of most community archives, let alone many large institutions, small organizations must be able to preserve their digital records in more flexible and cost-efficient ways. New work in the field, such as the interdisciplinary InterPARES Trust research project at the University of British Columbia, will also introduce further transformations related to how archives treat preservation and place. The Preservation as a Service of Trust project created by InterPARES is working toward a “set of preservation services that support the presumption of authenticity of records entrusted to the Internet” that can then be integrated as a standard within cloud storage environments. The ultimate hope is that by carefully defining and standardizing storage and preservation practices on the Internet, archivists may renew the trust in records preservation that has long been the core of our professional identity. Community archives have an equal role to play in this endeavour, and their unique circumstances and approaches to community-focused digital preservation may in turn inform and challenge current models and practices.

To close this article, I bring forth two different images, one old and one new, which speak to the promise of community archives regarding the preservation of their digital records. In 1982, Hugh Taylor made this incredibly prescient remark:

Not until television becomes interactive and is linked more closely with the computer will the new hybrid medium generate the full potential of all its parts and enable the individual and communities, through personally designed patterns of search, to explore and experience the past in the present which may be translated into active experience of ancestral stamping grounds or the local scene.

---


This description could easily apply to a set of networked community archives whose resources have been made available via the cloud: a “personally designed pattern of search” connects users to access copies of the digital records themselves acquired by the community archives and preserved via a trusted cloud service. More recently, in a 2014 advocacy paper, Lois Yorke describes the approximate number of Canadian archives as “900 points of light, large and small, stretching in chains and clusters and isolated dots across this vast land of ours.” It is no coincidence that Yorke describes the many community archives in Canada as dots of light: these archives are the data centres of our past. What remains now is to figure out how to connect the dots so that networked cloud solutions can continue to preserve the digital records of communities across nations.

Grant Hurley is a recent graduate of the Dual Master of Archival Studies and Master of Library and Information Studies program at the University of British Columbia. He also holds a Master of Arts in English from UBC. He currently serves as a director and editor of Off the Record for the Archives Association of Ontario and as a researcher with the InterPARES Trust project. He lives and works in Toronto.

71 Yorke, “Nine Hundred Points of Light,” 44.