Delayed while Pending Some Action: The Information Backlog as Deferred Maintenance

Ciaran B. Trace, School of Information, The University of Texas at Austin, cbtrace@austin.utexas.edu

“The task for scholars of infrastructure, then, is not merely to dwell on the materiality of infrastructure—an important topic on its own right, to be sure—but to dig deeper and uncover the humans who keep things going.”

(Fidler and Russell, 2018, p. 905)

Abstract

Repositories of historical records form part of the essential information infrastructure for humanities research. In the scholarly communications lifecycle, it is the archive (as place and as collection) that typically functions as the research laboratory, the source of knowledge for the humanities and its sub-disciplines. In this context, maintenance work involves upkeep of the network, the buildings, and the collections under care. In the archive, the collection level maintenance work of the archivist exists in an act of mediation between the creator and the user. Mediation is necessary because archival work is extractive. Documents exist as technical artifacts that are removed from their creating context. Archivists respond to this circumstance by engaging in restorative work - work to recreate that context for subsequent users of the archive. Embedded in this notion of the archive is the idea of a steady flow of information that resides within and moves through socio-technical systems. While there is exponential growth in the information transferred between the creator and the archive, the information flow between the archive and the user is often leaky – discontinuous and disrupted. In particular, there is a considerable interval between the time collections are accessioned, processed, and made accessible for research. It is in this space that what archivists euphemistically call ‘the backlog’ comes into existence. What the backlog interrupts is the distribution and consumption end of the knowledge process. As a concept and a reality, the backlog is a critical point of failure in this knowledge infrastructure, carrying with it an ongoing and prevailing sense that ‘deferred maintenance’ has become the norm in the archive.

Information Infrastructures and Knowledge Discovery

The past century has seen a boom in the development of a network of information infrastructures (II) in the United States that exist to support and sustain knowledge discovery. As Paul Edwards describes, these infrastructures exist as “robust networks of people, artifacts, and institutions that generate, share, and maintain specific knowledge about the human and natural worlds” (Edwards, 2010, p. 17).

In the social sciences, data archives emerged on university campuses from the 1940s onwards, providing preservation and access services for researchers looking to conduct longitudinal, comparative, and cross-sector research with quantitative data sets, at scale (Eschenfelder & Shankar, 2017). As an information infrastructure, data science archives rely on the work of data processors and data curators, who reformat and clean the data in an act of remediation as the data moves from the collection and analysis phase to one of subsequent reuse (Plantin, 2019).

On the other hand, repositories of historical records form part of the essential information infrastructure for humanities research. In the scholarly communications lifecycle, it is the archive (as place and as collection) that typically functions as the research laboratory, the source of knowledge for the humanities
and its sub-disciplines (Trace and Karadkar, 2017). It was historians who initiated the construction of the infrastructure of the archive in the United States, culminating in its modern form in the early twentieth century. Yet by the 1930s the still nascent historical profession had largely sloughed off this responsibility, with the occupation of the archivist emerging, in exile, as the labor to manage and maintain this infrastructure in public and private settings (Trace, 2017).

Figure 1. Photograph of the National Archives Building being constructed, 5/1/1934. ARC Identifier 7368457.

As a material base for research, archives contain documentary by-products of human activity – personal, organizational, and governmental. While the data archive is the primary home for surveys, enumerations, public opinion polls, and the like, the archive is the store-house of information and evidence about social life and social institutions that society deems valuable to preserve. The archive is where people go to “gather firsthand facts, data, and evidence from letters, reports, notes, memos, photographs, and other

1 The Society of American Archivists A Glossary of Archival and Records Terminology defines archives as “Materials created or received by a person, family, or organization, public or private, in the conduct of their affairs and preserved because of the enduring value contained in the information they contain or as evidence of the functions and responsibilities of their creator, especially those materials maintained using the principles of provenance, original order, and collective control; permanent records.”

2 With a history of over 200 years of development, archival repositories in the United States are run by (federal, state, local, and municipal) governments, universities, for profit and non-profit organizations, libraries, and historical organizations.
primary sources.”

The notion of ‘the archive’ embodies that part of the American psyche that seeks to connect with history and culture, to strengthen community and a sense of belonging, and to protect people’s rights and entitlements.

As an infrastructure created by a professional community, archives exist as standalone, networked, or co-located entities whose workers harbor a sense of commonality through professional standards and collaboration, and membership in professional bodies. As an infrastructure created by discreet organizations, archives exist as institutional arrangements, represented by particular physical and organizational structures and facilities. As an infrastructure created by artifacts archives exist as secure information systems - conduits from the creator of the source to the subsequent user of the source, with the collections existing in permanent escrow with and within the archive.

In all these notions of ‘the archive as infrastructure’ temporal dimensions loom large. The archive (in its various manifestations) has set itself up to be the preserver of the historical record – ‘forever,’ ‘in perpetuity,’ or in the case of the National Archives, ‘for the life of the republic’ – take your pick.

**Information Infrastructures and Maintenance Work**

The combination of infrastructure and time inevitably lead to questions of upkeep. Upkeep is a state of being, an act, a process, a means to an end. In the research literature ‘upkeep’ is synonymous with the idea of ‘maintenance.’ In academic disciplines and professional practice, maintenance is a lens (“a theoretical framework, an ethos, a methodology, and a political cause”) that centers our experience of a world in which the breakdown and repair of infrastructures (political, civic, physical, social, etc.) are commonplace (Mattern, 2018). As Edwards et al. remind us “Infrastructure is all about maintenance. Maintenance, maintenance, and more maintenance. It doesn’t just get built, like some colossal monument left to stand until natural forces wear it away. It constantly has to be repaired, rebuilt, extended, shrunk, adapted, readapted, continually redefined and reengaged” (Edwards et al., 2011, p. 1409).

In their article about the history of maintenance work, Russell and Vinsel ask us to ponder “What things persist over time, and why?” (2018, p. 7). In the example of the archive, maintenance work involves upkeep of the network, the buildings, and the collections under care. But it did not take two centuries for the concept of maintenance to seep into the archival imaginary. The story of the archive in the United States has been one of ongoing conditions and acts of maintenance and of repair.

Channeling Denis and Pontille (2015), the infrastructure of the archive can be understood in terms of its various modes of existence. In pursuit of legitimacy, the infrastructure of the archive has been driven toward stabilization, manifested in a concern for standardization and consistency of work process and product. Yet, from the outset, the infrastructure of the archive has also been characterized by its fragility. The notion of fragility is tied to the material and concrete reality of maintaining an infrastructure whose bedrock principle is permanence. Similarly, the notion of fragility “describes a component of infrastructure that is subject to failure or degradation, usually due to uncertain availability of the resources necessary to sustain it” (Borgman et. al, 2016, p. 1). In this instance, the continual struggle for resources is one result of the lack of social resonance to that which the archive clings - to the past, to history, to the study of human culture, and the like. This is an infrastructure that has long been subjected to political buffeting as

---

society struggles to characterize the archive as “an administrative, economic, or social good” (Trace, 2017, p. 134).

Stories of fragility abound within the network of archival infrastructure. Such is the case of the Georgia state archive where the cyclical political and social forces of administrative and bureaucratic reform have long served to undermine the survival of this state agency, including in the first few decades of the twentieth century (Trace, 2017). An unappreciated role taken up by the first female state archivist of the State of Georgia (1925-1937), Ruth Blair, was infrastructural care work. Work in which she strove to ‘keep going,’ ‘patch up,’ and ‘set right’ parts of the state’s archival infrastructure that were continually under threat.

![Figure 2. View of Rhodes Hall, 1905. Image from the Kenan Research Center, Atlanta Historical Center.](image)

During her tenancy, Ruth used her personal connections to secure the removal of the Archives Department from the cramped conditions at the State Capitol to Rhodes Hall, a Romanesque revival style twenty-two room mansion on Peachtree Street in Atlanta, formally known as ‘Le Rêve’ (see figure 2). While the legislature appropriated $5,000 to repair the building and install shelving, the economic realities of depression-era Georgia provided cover for a situation in which no state money was appropriated for the ongoing maintenance of the building, an expenditure estimated to cost $1,500 annually ($750 for the porter, and a similar sum for the cost of heat, light, water, and incidentals). With the state abrogating its duty as maintainer, Ruth responded with commitments to sustain the archival infrastructure. As part of her work of caring for the archive, she paid for the upkeep of the building, an undertaking that cost her at least $1,800 before the situation was resolved (Trace, 2015).

**Information Maintenance and Repair**
Yet, it is not just the external infrastructure of the archive that must be maintained for the research process to function. Similar to Barley and Bechky’s study of technicians in scientific labs, archivists, as technicians for the humanities, manage “an interface between a larger work process and the materials on which the process depends” (Barley and Bechky, 1994, p. 88). In the language of goods and services, the archive exists as an infrastructure of information production, distribution, and consumption. Embedded in the notion of the archive is the idea of a steady flow of information that resides within and moves through socio-technical systems - from the creator to the archive to the researcher.

According to Downey, information labor exists to “set information in context, to move information across context, and to reset that information in a new context” (Downey, 2014, p. 149). In the archive, the collection level maintenance work of the archivist exists in an act of mediation, or what Barley and Bechky (1994) call ‘brokerage,’ between the creator and the subsequent user. Mediation is necessary because archival work is extractive. Documents exist as technical artifacts that are removed from their creating context. Archivists respond to this circumstance by engaging in restorative work - work to recreate that context for subsequent users of the archive. If the goal of social science data curators is that the artifacts be “pristine” in order to be reused (Plantin, 2019), the goal of the archivist is to resettle the artifacts as evidence in place, but to do so within the confines and limitations of the archive (Trace and Francisco-Revilla, 2015).

An elemental concern for setting information in context means that archivists strive to maintain the link between the records, their creators, the functions and activities that brought the records into being, and the recordkeeping structures in which they were originally stored. In effect, the goal is to fix the original socio-technical context in place through the physical and symbolic work that takes place within the archive. This maintenance work is methodologized through guiding principles and conventions of practice. Birthed in the ideas of the profession’s European antecedents, American archivists draw from two fundamental principles in undertaking maintenance work on collections. The principle of respect des fonds, dictates that “records of different origins (provenance) be kept separate to preserve their context.” The principle of original order dictates that “the organization and sequence of records established by the creator of the records” must be maintained (SAA Glossary).

This maintenance work is given expression in the archive through the act of ‘archival processing.’ Processing allows the archivist to gain intellectual and physical control of archival materials through the activities of accessioning, arrangement, description, and preservation. In the analogue world there is a distinct physicality to processing, although this work is obscured or rendered invisible from public view, happening as it does in the private staff areas of the archive.

In this private space, processing acts as a ‘concealed tool’ of maintenance and repair (Graham & Thrift, 2007). As part of this bench work, archivists physically inspect the records, carrying out preservation and conservation treatments to minimize deterioration of the materials (and the information they carry) and maximize their projected life span. Archivists also arrange or rearrange archival materials into recognized groupings (the evidentiary layers of record group, subgroup, series, and files), while also inscribing context in a descriptive finding aid for the collection. In these ways evidence is stabilized, set in place physically and intellectually, ready for the user to discover.

The notion of repair is ever present here. Creators’ original recordkeeping structures atrophy when removed from their original context, and materials themselves succumb to the elements and degrade
over time. Thus, archival maintenance also “occupies” and “constitutes” what Jackson calls “an aftermath” (Jackson, 2014). The repair in this instance, is an attempt to heal the rift in space and time that has opened up between the materials as created and the materials as archived, an attempt to restore the materials to a place in time and to a condition where they have the most meaning.

![Figure 3](image1.png)

Figure 3. Frank B. Crawford reinforcing a document with crepeline in the [National Archives] Division of Repair and Preservation, 1940. ARC Identifier 3493246.

![Figure 4](image2.png)

Figure 4. Photograph archivist Sara Jackson at her desk in a stack area of the National Archives Building, 1935, ARC Identifier 122213584.

Throughout this physical and intellectual process, all traces of the archivist explicitly and implicitly recede from view. The rhetorical strategy of the written finding aid, for example, works to erase the agency and emotion of the archivist from the narrative (Trace and Dillon, 2012), relegating the profession to a subordinate role that Cook has equated to that of “an invisible caretaker, a docile handmaiden” (Cook,
2010, p. 3). Despite or perhaps because of its invisibility to the outside world, the nature and composition of the labor that undergirds maintenance work has recently been the focus of intense scrutiny among archival practitioners. Echoing Jackson, the profession has come to understand that “questions of visibility and invisibility may be intimately linked to power” (Jackson, 2014, p. 229). That this labor is often contingent, done by workers on temporary contracts, highlights one set of concerns about the ethics and sustainability of the maintenance process. That, typical of other feminized professions, the emotional labor inherent in such work goes unrecognized, has been an additional source of tension.

Although these concerns are deserving of further research, this paper concludes by examining another notion of visibility and invisibility as it plays out in infrastructural and maintenance work. Here I refer to the notion that both the infrastructure and associated maintenance work are often hidden from view until a breakdown or failure occurs (Star, 1999). In the archival world this breakdown, this failure to maintain, is manifested in the ever-present ‘backlog’ that haunts most archival institutions. To paraphrase Knowles (2016), the backlog can be understood as part of a long process of degradation and deferred maintenance on (analogue and technological) systems built to support knowledge infrastructures.

**Failure, Deferred Maintenance, and the Notion of the Backlog**

In the archival pipeline, the rate of flow between the creator and the archive continues to grow exponentially (Greene and Meissner, 2005). Yet the information flow between the archive and the user is leaky – discontinuous and disrupted. In particular, there is often a considerable interval between the time collections are accessioned, processed, and made accessible to the public. It is in this space that what archivists euphemistically call ‘the backlog’ comes into existence. In the professional literature, the backlog is defined as “materials received by a repository, but not yet processed… Anything delayed while pending some action” (SAA Glossary). Comprehensive statistics on the size and the growth rate of the backlog in the United States are hard to ascertain, but there are abundant examples from which to draw.

In the mid-2000s, for example, the US National Archive and Records Administration estimated their backlog at 1 million cubic feet of records. An analysis of their textual holdings found that “74 percent of textual records in Washington were not processed sufficiently to enable researchers to easily identify records of interest” (Bucciferro, 2008).\(^4\) At the end of fiscal year 2012 NARA had made some headway, with the processing backlog sitting at 40 percent. However, in the same year over 95 percent of electronic records held by presidential libraries were believed to remain unprocessed (NARA Office of Inspector General, 2013). As of June 2018, the Texas State Library and Archives Commission had a backlog of about 32,900 cubic feet of state records, almost 41 percent of the total holdings (Sunset Advisory Commission, 2019).

What the backlog interrupts is the distribution and consumption end of the knowledge process. In essence it reflects an interruption of the information flow leading to an associated service disruption. As a concept and a reality, therefore, the backlog exists as a critical point of failure in this knowledge infrastructure. It carries with it an ongoing and prevailing sense that ‘deferred maintenance’ has become too often the norm in the archive (Scott Gabriel Knowles, 2016).

---

\(^4\) Thirty-three percent of NARA records did not possess basic elements of intellectual control, such as titles or dates.
Figure 5. Photograph of U.S. Food Administration Records Arriving at the National Archives Building, 1935, ARC Identifier 7820503.

But whose failure is it? Graham and Thrift state that “attention to the need for repair and maintenance of infrastructures tends only to occur after catastrophic, rather than prosaic failures” (Graham & Thrift, 2007, p. 9). One could argue that the archival backlog is, in fact, that ignored and prosaic failure that has played out gradually and cumulatively across time. In this respect, the failure is a result of a ‘slow disaster’ (Knowles, 2016) that has manifested itself as a dearth of funding and an associated lack of investment and care in the cultural heritage sector. That the infrastructure is being undermined from without should come as no surprise. After all, as Edwards et al. remind us, information and knowledge infrastructures are “embedded in and over laid across cultures, organizations, governments, and other social forms,” forms that have the power to not only interpret infrastructures but to erode it (Edwards et al., 2011, P. 1409).

But what then is the attitude of the archival profession to the eroding of its infrastructure? According to Edgerton, maintenance lives in a “twilight world, hardly visible in the formal accounts societies make of themselves” (Edgerton, 2011, p. 79). In the world of the archive, the attitude toward this deferred maintenance is diffuse. On the one hand, the notion of deferring maintenance on collections has long been normalized, if only internally. The term ‘backlog’ forms part of the lexicon in archival glossaries and manuals and has recently been reified in the architecture of digital preservation systems. But the backlog has often been hidden from external view, with the presence of unprocessed collections elided from catalogs and thus from the users of the archive.

At times, however, deferred maintenance has explicitly been made visible to external constituents, turned outward in a bid for institutional and professional legitimacy and sustainability. Archivists have sought their way into the public consciousness by declaring the presence of the backlog. In pursuit of resources to rectify the issue, archivists have drawn attention to so called ‘archival silences’ - the silences that ensue within the historical record when collections are unprocessed and thus unavailable to researchers. When

---

5 See, for example, the ‘backlog’ tab in the open source digital preservation system, Archivematica. https://www.archivematica.org/en/
these appeals are successful, the resulting funding from granting agencies has resulted in the creation of special, usually temporary, positions to help address the issue.6

The archival profession has also used the crisis of deferred maintenance as a catalyst for change, in particular seeking to drive efficiency and improvisation in practice. The 2005 article, “More Product, Less Process,” by archivists Greene and Meissner (2005), was a jolt to the profession. With its call to fundamentally reengineer archival systems for greater maintainability, the authors link the backlog not to the systemic failure or constraints of social and political systems, but to a failure on the part of the maintainers – a failure the authors believe is manifested in archivists’ “professional fastidiousness,” “excessive pride in craft,” “reluctance to be perceived as sloppy or uncaring by users,” and mere concern with “housekeeping issues” (Greene and Meissner, 2005, pp. 233-234, 241).

In Greene and Meissner’s argument the shock comes from the notion that the enemy is us, the claim that the infrastructure is being undermined from within. The failure in this scenario is not just a failure to maintain, but a failure on the part of archivists to innovate when faced with an influx of large twentieth-century collections. Self-care in this scenario involves prioritizing funders and users of the archive and doing so with no increase in resources but with a calculated change to the nature and speed of maintenance work. That this particular call for self-healing constitutes an implicit embrace of a neo-liberal agenda for information labor and maintenance work, an agenda that has already served to hasten the infrastructure into disrepair, has not gone unnoticed or unchallenged (Cifor and Lee, 2017). Indeed, the conflict in the story of the backlog highlights the fact that the archive is a contested space, a space where competing narratives emerge of what it means to maintain and invest in infrastructure.

References


6 See, for example, the Cataloging Hidden Special Collections and Archives Program (2008-2014) from the Council on Library and Information Resources (CLIR).


