 FEATURE: GLOBAL PERSPECTIVES

E-Mail as Legacy: Managing and Preserving E-Mail as a Collection

Jesse David Dinneen and Maja Krtalić

abstract: This article discusses how e-mail and e-mail archives can serve as legacy for individuals, families, and societies. Toward this end, it outlines the scope and scale of e-mail as a tool and a task, and gives an overview of the many individual, group, and institutional factors in e-mail management that influence the form and contents of e-mail archives. Finally, drawing on the authors’ experiences and observations as personal information management (PIM) researchers, the article discusses e-mail management practices, makes recommendations for effective e-mail management, and suggests future research topics that span library and information studies.

Introduction

In today’s information society, e-mail is ubiquitous as both a communication technology and a task: most knowledge workers must not only read and write e-mail every day but also decide whether to save or delete those messages and how to organize them or not. As e-mails are kept and e-mail archives produced, such communications increasingly become a resource. E-mail archives can serve as legacy for individuals, recording weddings, births, and other milestone events within a family, and for society as collective memory and national heritage.

E-mail today has been appropriated (or overloaded) by users for functions well beyond the basics of communication and messaging. People use e-mail in various ways for coordination and collaboration, record-keeping, personal information management (PIM), and integration with other technologies, such as files. E-mail has also grown considerably in scale, with global e-mail traffic likely surpassing 300 billion messages per day. Individual e-mail archives

Copyright © 2020 by Johns Hopkins University Press, Baltimore, MD 21218.
(as collections) have grown accordingly, increasing roughly 10 times in size from 1996 to 2006. Despite, or perhaps because of, the scale and complexity of e-mail use, challenges in even the basic use of this messaging system are common, and systems and interfaces evolve slowly to address such issues.

The story is no less complex once e-mails are aggregated into an archive, since the form and use of such archives are determined by multiple stakeholders and end users, including contributors (for example, author-senders, reader-receivers, donors, and citizen scientists), curators, archivists, librarians, researchers, and consumers. As this article discusses later, many factors influence each stakeholder and therefore the e-mail archive.

The following sections consider the idea of e-mail as legacy, beginning with an outline of the factors that influence e-mail and thus e-mail archives at individual, group, and institutional levels. After discussing how e-mail can constitute or contribute to different kinds of legacy, the article recommends principles for effective e-mail management at various levels and suggests important areas of research across library and information studies. Throughout, the authors draw on both existing literature and their experiences and observations as PIM researchers.

Factors Influencing E-Mail Management

Individual Factors

The factors influencing how individuals manage e-mail are important to understanding e-mail archives because most records are initially made—read, written, stored, and organized—message by message, by individuals. In other words, factors influencing PIM behavior determine the shape of e-mail archives. Most time is spent reading and writing, little is devoted to organizing, and different strategies have emerged to keep e-mails in order. For example, people may be prioritizers, archivers, no-filers, spring cleaners, frequent filers, or folder cleaners. Which strategy is best will differ among individuals, tasks, and other circumstances. A system will likely arise in response to external factors, such as users’ occupations and assignments; the types of information sent and received; the message transmission rates and overall volume; the collection size, organization, and age; and software and interface functionality. The choice of strategy is also affected by internal considerations, such as users’ preferences (for example, search or navigation), cognitive abilities, desire for order or an empty in-box, familiarity with the collection, risk aversion, and hoarding habits.

Group Factors

Group management of e-mails is determined by individuals, their sharing habits, the mutual task, the software and software ecosystem used, the software’s integration with other tools, and each user’s implicit and explicit expectations. The result is as complex as any undertaking involving human interaction. Such activity requires the involved
parties to identify (or create) and use, implicitly or explicitly, strategies for organizing knowledge and for doing so collaboratively. Many projects involving group-managed e-mail result in reinventing the wheel of “knowledge organization,” the process of capturing, distributing, and effectively using knowledge, which is familiar to information professionals and has been a research topic for decades. The usually implicit negotiation of group management is further complicated by mutual intelligibility. Personalization of meaning may conflict with its generalization, so the more an e-mail archive makes sense to one person (for example, by having folder names and organization that seem reasonable to that individual), the less logical it may seem to others.

Institutional Factors

E-mail management becomes further complicated when considered from an institutional or organizational perspective. In an organization, what e-mails should be managed, and why, by whom, how, and when?

The broadest institutional factors determining the management of e-mail and e-mail archives are those regarding the value of the messages. In other words, the “why” question addresses the purpose and context of e-mail management (why should e-mails be managed?), while the “what” question addresses e-mails’ content and value (what is worthwhile and should therefore be managed?). E-mail serves as a communication channel, but it is also a product of organizations’ business activities and as such is often a record containing authentic and evidentiary information and supporting business continuity. Through e-mail, organizations may set or interpret policy, formalize business processes, or provide evidence of the activities of an office or department.

Through e-mail, organizations may set or interpret policy, formalize business processes, or provide evidence of the activities of an office or department.

The purpose and strategies of managing e-mails as documents are well addressed in the literature of records management. E-mails can also serve as a knowledge base and as a memory of an institution, tracing and recording decision-making processes, management style, and the context of those decisions. As such, e-mail is a valuable tool
for evaluating not just what an organization did and when but also how it did so and why. Research addresses this issue by evaluating existing techniques and developing new ones, for example, by automatically sorting e-mails by similarity of content and context and connecting them with expressions of organizational knowledge.\textsuperscript{14}

If organizations can clearly state the purpose of e-mail (historically, presently, and in the future), and how e-mail fits into the wider context in which the organization operates, doing so will allow deeper reflection on the content and value of e-mails. Further, everything that happens in workplaces is a part of wider social, cultural, and legislative milieu and is subject to constant technological change, which in turn affects behavior and expectations in work environments.\textsuperscript{15} E-mails are therefore also reflections of the community and social relationships present in the organization. For all these reasons, e-mails clearly have primary and secondary—and explicit and implicit—value.

Narrower factors in institutional e-mail management surround logistics and practicality (that is, who, how, and when). The “who” question addresses people and their roles and skills in managing e-mails. Most e-mails are written by individuals, but sometimes generated automatically, so who should make sure e-mails are accessible, retrievable when needed, usable, and understandable? Further, who is responsible for ensuring the accurate, authentic, and reliable representation of e-mail transactions and content? While answers to these questions will differ across organizations and over time, a sense of shared responsibility and clear and guided practices must pervade the answers. If everyone involved in the process of managing e-mail understands their role and contribution and attempts to follow organization-recommended practices, e-mail preservation and management will more likely succeed.

The “how” and “when” questions address the fine-grained processes of e-mail management and archiving: how (that is, with what processes or guidelines) and when (or until when) should e-mails be managed, archived, and retained? These questions indicate important institutional factors in e-mail management, but as with the “who” question, concrete answers will depend on the organization, and this article cannot provide them. Fortunately, many prior works provide practical, managerial, and technical advice about e-mail archiving processes.\textsuperscript{16}

\begin{center}
\textbf{If everyone involved in the process of managing e-mail understands their role and contribution and attempts to follow organization-recommended practices, e-mail preservation and management will more likely succeed.}
\end{center}

E-Mail as Legacy

Many people nowadays write e-mails daily. In doing so, they capture traces of their lives and work, for example, through correspondence with friends and family about major life events or by storing evidence of professional development. These traces can be valuable sources of information about facts, events, social atmosphere or milieu, and culture at a place and time. Documentary and evidentiary value contained in e-mails can impact
individuals themselves, their family members, organizations where they might have worked, community archives to which they may have contributed, or institutions such as those in the heritage sector that might collect certain categories of e-mail. Because of the documentary and evidentiary values, preserving e-mail can be of interest to society writ large. In other words, an e-mail can contribute to an individual’s information legacy, and thus for some individuals, to a society’s heritage as collective memory. But what kind of legacy does it provide, and what is the value handed down to future generations—that is, what is the heritage value? To answer these questions, consider some of the features and context of e-mail through the lenses of personal information management and legacy generally.

E-Mails in Personal Collections

Many people, especially later in their lives, are faced with challenging decisions on whether and how to preserve their personal collections of documents and artifacts, including e-mails, and how to tell their life story. If those decisions are not made by the people who own the materials while they still can, decisions fall upon their family. After loved ones die, family members must deal with a range of organizational, emotional, and sometimes legal issues that impact what they do with inherited personal collections. Some examples of those decisions are

- choosing what is relevant from the lifetime collections of things,
- understanding and extracting hidden data that have value for family history or legacy,
- maintaining access to collections,
- deciding how to share collection items with other family members or the public, and
- addressing copyright and privacy issues.

Even though plenty of advice is available on how to organize personal collections, especially in the digital environment, these decisions can be so overwhelming for many people that the “delete all” mindset may seem the only way out, despite the loss of value it entails.

Personal Collections of Significant Individuals

When it comes to preserving e-mail as a legacy for society, one category of people could be particularly relevant: significant individuals, who have created something valuable for society in their life and work. People working in the creative and public sectors—writers, artists, public servants, and the like—are particularly interesting because their personal collections contain a wealth of valuable information about their work, their society, and that society’s history. Such materials therefore have high potential for future reuse and cultural value. After all, much research has been done on handwritten letters preserved in libraries, archives, and museums, and arts and humanities would look different today if those rich sources of information had not been available.

Numerous curation and preservation issues must be considered and addressed in the transition from personal to joint collections in libraries and other institutions. Some of
those issues are documenting and understanding context, managing privacy and sensitive information, and handling copyright issues in access and use. These problems have lain at the heart of research for many years, and the library literature gives guidance on how to curate and preserve digital collections.\textsuperscript{19} Successful transition from personal to collective domains, however, relies on the fact (or assumption) that individuals have indeed preserved and described their collections, whether because of awareness of the legacy potential or for some other reason, such as a personal need for organization. E-mails are an essential component in those collections. Initiatives such as Stanford University’s open source software ePADD have been developed to support archival processes for e-mail collections.\textsuperscript{20} The letters in the name ePADD refer to electronic processing, appraisal, discovery, and delivery. Therefore, collaboration between individuals and heritage organizations such as libraries is one key to successful acquisition and subsequent management of personal collections.

Recommendations for Effective E-mail Management

Recommendations for Individuals

Individuals should seek out, discuss, and try different PIM strategies and systems. As mentioned, the best organizational strategy will differ among people and tasks, and even across time for a single person. Therefore, occasionally considering differing strategies is important. You may not find a more effective plan, but you will not hit upon one if you fail to look for one or to reflect on the task. Some strategies are simple. If you receive a high volume of e-mails, for example, try triaging messages of low priority first to avoid feeling overwhelmed by the sheer number of communications and then address those that remain. People with personalized, committed strategies report higher levels of satisfaction with e-mail use and management.\textsuperscript{21} One specific aspect of personal e-mail management to consider is the trade-off between time invested organizing, such as creating tags and folders, and time spent retrieving, such as searching or looking through tags and folders. If you seldom need to find old e-mails or find it easy to do so, it may be better to save time initially by not organizing messages, and to accept that they may take a little longer to retrieve later. Active organizing and occasional cleanups take time, but searching by keywords or even dates tends to be effective in retrieving e-mail, and storage is usually abundant.\textsuperscript{22}
however, that this approach precludes adding useful metadata and makes an archive harder for family members, curators, and others to reuse later. Conversely, if you find retrieving e-mails difficult, you may benefit from investing more time in sorting messages into folders or adding tags soon after receipt, processes that can be aided with automation.23 Later, the folders and tags can be used—by you and by others—to refind your messages and make sense of the collection.

**Recommendations for Groups**

As with PIM prescriptions for individuals, evidence-based prescriptions for group management of information are currently minimal, in part because the topic is still gaining attention and focus in research.24 However, a few actions can ease the group management of e-mails. A relatively simple first step is to acknowledge as a group that you are collaboratively organizing information, and thus explicate conventions for the use and naming of folders and tags, and clarify expectations around version control, duplication, retention, and so on. Members should be encouraged to think of the other users or to think of the next user (including themselves, later in time) and future use of the information. Such awareness will aid knowledge continuity and avoid loss of information (and precious work). For these reasons, it may also be useful to adopt and assign position-specific e-mail addresses so that when one member exits a project, information accumulated by that person about the project does not disappear too. If the whole archive cannot be saved, metadata could be retained. But remember that the more convenient people find the guidelines, the more likely they will follow them, while inconvenient rules may be ignored.25

**Recommendations for Organizations**

Organizations have access to extensive advice published in professional literature and other forms of sharing good practice. A few simple recommendations, however, will address the issues discussed in this article.

*Have Clear, Accessible, and Up-to-Date Policies*

While many organizations recognize the benefits of having information and records management (RM) policy or strategy, their policies are not always easily findable, accessible, or comprehensible. When personal factors mediate the interpretation and implementation of such policies, different practices might occur for even a simple instruction. To avoid misunderstanding and to minimize the diversity of personal attempts in managing documents and records, including e-mails, the first step is to make sure that the guidelines are clear and relevant to employees’ real practice. For example, consider the format and

Training that results in skills applicable in personal life as well as in the workplace will be far more efficient and will enhance understanding of the value and purpose behind preserving e-mails.
content of such policies. Written formal policies are only one way to present the content; instruction in the form of videos, storytelling, decision trees, and the like can reach people more directly. Training that results in skills applicable in personal life as well as in the workplace will be far more efficient and will enhance understanding of the value and purpose behind preserving e-mails.26

Have Procedures for Common Tasks, but Factor in the Personal

It is helpful to have automated processes in place for such common tasks as naming and adding metadata and to use special software to reconcile authority records, make finding aids, remove confidential information, aid browsing, auto-create tags and folders, and the like. In addition to automatic processes, however, organizations should factor in the personal. For example, people sometimes tend to transfer usual practice from one workplace to another or look for ways to simplify processes. Another example of factoring in the personal is what happens to e-mail collections when people leave an organization. The e-mails may become lost, inaccessible, or practically unusable, resulting in the loss of valuable knowledge and evidence. Especially when people retire after many years of service, their e-mail and document collections might contain personal insights and help bridge organizational knowledge gaps.27

Consider Sustainability

Sustainability can be considered from several perspectives, such as technological, environmental, cultural, and societal, and this article cannot address all of them. Managing disruptive events—from staffing changes to digital and natural disasters—is crucial to organizational sustainability. E-mails can contribute to the stability of a business or other organization by ensuring access to authentic and reliable data, such as demonstrating compliance with legislation. Sustainability issues therefore require having awareness, policies, and protocols in place to deal with technologies used to curate and archive e-mails, and to manage interoperability and interconnectivity of e-mail management systems. Of course, every digital solution for handling e-mail impacts the environment to some extent. For example, e-mail storage requires electrical power, still primarily produced by burning fossil fuels. Further, e-mails managed within every organization are embedded in the national culture; cultural sustainability therefore reflects in protocols around access, interpretation, and use and reuse of e-mails.

Among basic sustainability practices, considering “so what” and “what if” scenarios is powerful for raising awareness and addressing critical issues. Sometimes, the implications of e-mail management are not obvious until a situation happens, whether positive or negative. At such times, access to e-mail, its content, and its context becomes critical for organizations’ continuity (for example, during a pandemic that prevents physical access to a workplace). Considering such scenarios can thus help foresee what could happen and how organization-prescribed procedures may affect the outcomes.
New Zealand Perspectives

There are many institutional initiatives to preserve e-mails in organizations in New Zealand, either as business records or as part of heritage collections. In January 2020, the National Library of New Zealand held a symposium bringing together archivists, librarians, researchers, records managers, and others from the cultural heritage and information management sectors. The Archiving Email Symposium was a place to hear about issues and challenges of preserving e-mail archives in government departments, local councils, archives, and academic and research libraries, and in the personal context. Participants shared similar experiences and discussed challenges around legislation; standards; software and tools; processes, such as collecting and managing e-mail at scale, transfer, and ingest; preservation formats; and access and use of e-mail archives. Hearing about practices in diverse contexts and sharing stories of both success and failure was a significant contribution of this symposium. It also confirmed that the boundaries between individual, organizational, and societal need not be seen as borders but rather as pathways to more efficient e-mail preservation.

Conclusion

E-mail and e-mail archives are complex phenomena, in part because of the many stakeholders involved and the many factors influencing actions at each level. E-mail’s persistence and use across time allows it to function as a legacy for individuals and society, which adds significant value but also complications to e-mail management that require careful consideration and further research. It is important to consider and account for the personal factors involved in e-mail management and to study e-mail archives from multiple perspectives, not only from those of subfields, such as PIM and RM, but also from those of other research interests. The considerations overlap with those of personal (digital) archives, information behavior, information retrieval, digital libraries, knowledge organization, and metadata. Further, improving practices surrounding e-mail archives requires multiple kinds of practical solutions at different levels: more and better data and data collection tools, clearer policy, and improved user interfaces, to name just a few, and has implications for how libraries and librarians acquire and manage electronic personal archives in general. Researchers should therefore directly draw on and integrate the appropriate subfields in studying the handling of e-mail as a complex and ubiquitous information management task.

Jesse David Dinneen is a senior lecturer in information systems in the School of Information Management at Victoria University of Wellington in New Zealand; he may be reached by e-mail at: jesse.dinneen@vuw.ac.nz.

Maja Krtalić is a senior lecturer in information studies in the School of Information Management at Victoria University of Wellington in New Zealand; she may be reached by e-mail at: maja.krtalic@vuw.ac.nz.
Notes


17. Evan Carroll and John Romano, Your Digital Afterlife: When Facebook, Flickr and Twitter Are Your Estate, What’s Your Legacy? (Sebastopol, CA: Safari, 2010).


