

Personal Email Management on the University Digital Desktop: User Behaviors vs. Archival Best Practices

Megan A. Winget

University of Texas at Austin; 1 University Station, D7000, Austin, TX

78712-0390 megan@ischool.utexas.edu

Kimberly Chang

University of North Carolina at Chapel Hill, School of Information and Library Science, 100 Manning Hall, CB 3360; Chapel Hill NC

27599-3360 kimchang23@yahoo.com

Helen Tibbo

University of North Carolina at Chapel Hill, School of Information and Library Science, 100 Manning Hall, CB 3360; Chapel Hill NC

27599-3360 tibbo@ils.unc.edu

This paper will review findings from an extensive user study which seeks to understand the ways in which individuals currently manage, interact with, and think about electronic files, particularly email. Funded by the National Historical Publications and Records Commission (NHPRC), we web-surveyed nearly 3000 faculty, staff, and administrators; and conducted personal follow-up interviews with 100 people across two universities: one public, one private. Our findings indicate that users and archivists have competing concerns. User concerns revolve around the volume of files with which they are forced to contend, leading to wasted time, improperly deleted files, confusion, and general dissatisfaction with the networked environment. Archivists, on the other hand, concerned with the preservation of digital materials, are focused on organizational issues that are often at odds with users' daily practice.

Introduction

Networked computing has become an omnipresent feature of everyday life. Nominally a 'good thing,' this kind of access allows users to create, gather, transmit, and interact with

vast amounts of varied data in a few keystrokes; but the technology lacks a mechanism to help users make sense of all the information. Numerous studies have suggested as much. Barreau and Nardi (1995) have said that although useful tools have been developed for handling archived information, users have a great deal of difficulty maintaining access to large quantities of ephemeral information. Whittaker & Sidner (1996) have described email use - its character as a catch-all repository for both exceptionally important and remarkably mundane information. In a cross-tool longitudinal study related to Personal Information Management, Boardman & Sasse (2004) report that users employ multiple strategies to manage different types of information, and the design of technological tools must take such variations into account. While there are numerous reports detailing personal information management habits and the use of technological tools to manage information, there are no studies the authors could identify that compare these habits and the tools with accepted archival and records management practices.

The study for which we are offering this report seeks to describe in greater detail the native behaviors of users in an academic setting and discover how to optimize such behaviors to serve standard archival and records management requirements. The project is a collaborative effort between academic professionals and practicing archivists at two major research universities - one a private institution, the other a public. This three-year project is examining the desktop computing environments and behaviors of faculty, administrators, and university staff at these two universities, in an attempt to understand how individuals interact with their data: both electronic files and email.

For many users, managing the vast amount of data on our computers is monotonous, time-consuming, and frequently unsuccessful. Although it's easy to create digital documents and communicate via email, our existing software systems give little direction on how to name, organize, or retrieve these materials. Not only is this situation challenging and frustrating for the user, but it also makes records management and the subsequent archiving of materials in electronic formats difficult, if not impossible. Despite the considerable amount of research conducted on personal information management systems, there is little known about these behaviors in an academic setting and the direct effect of such behaviors on records management programs.

In the university environment, because of financial and legal constraints, individual information management behaviors determine the level of an institution's records management and archiving success. The Managing the Digital University Desktop project, a collaborative effort between academic professionals and practicing archivists, is a three-year project seeking to examine these management practices. Funded by the National Historical Publications and Records Commission, the project team studied electronic file management behaviors by web surveying nearly 3000 users at both public

and private universities, then conducting follow-up interviews with 100 faculty and staff members at the two institutions. The interviews were centered on questions regarding the appraisal, arrangement, authenticity, physical preservation and privacy of electronic records and email. One of the project's goals was to describe, in greater detail, the native behaviors of users in academic settings, and discover ways to optimize their digital behaviors in order to more effectively serve standard archival and records management requirements.

Users' email attitudes and practices were specifically interesting to our team. In particular, we were seeking the context-sensitive data and user insights necessary for the successful future implementation of an Electronic Records Management System (ERMS).

Methodology

Understanding user needs and behaviors is essential to improving any computer human interaction. Gould, Boies, and Lewis (Gould, Boies, & Lewis, 1991) detail a usability design process that involves four elements: early focus on users; integrated design; early and continued user testing; and iterative design. They point out that "several lines of evidence indicate that this usability design process leads to systems, applications, and products that are easy to learn, contain the right functions, are well liked, and safe." Further, Thomas Sheridan (1989), in *Technological Forecasting and Social Change* argues that design of complex technologies should be seen as being "of, by, and for the people."

In September 2002 we distributed an electronic survey regarding basic email behaviors to the entire faculty and staff populations at two universities: Duke University (hereafter referred to as Duke), and the University of North Carolina at Chapel Hill (hereafter referred to as UNC). We sent announcements of the project and invitations to participate to all faculty, administrators and staff who were accessible via the universities' mass-emailing systems. The survey was sent to approximately 25,800 registered email addresses. (As some individuals had more than one address at one of the universities, the total number of individuals contacted was somewhat less.) We followed this invitation a few days later with a message asking for participation either through a web form or by returning an attached survey via email. Discounting any responses that we could not link to specific people, we closed the response period in October 2002 with 2,936 valid responses across the two campuses. Participants came from a wide variety of departments and fields, and included representatives from all job levels, ranging from staff, faculty, and higher-level administrators.

The campus-wide survey was designed to assess daily email and file use and discover such information as the type of software being used, the number of emails received per day, and the amount of time spent on email. According to the survey findings, 53% of

those surveyed at UNC and 38% at Duke received 1-20 emails per day. Duke staff reported spending the most amount of time on their email, with an average of 111.50 minutes per day, followed by Duke faculty with 92.69 minutes per day. UNC faculty and staff both reported spending on average approximately 81 minutes per day on email.

The survey revealed concerns on both campuses about the content, access, and management of email and electronic files. The top concern listed was unsolicited email in the form of advertisements, spam, or other junk mail; followed closely by confidentiality - the concept that their messages may be accessible to individuals the sender or receiver does not intend. While respondents from both campus complained about the lack of time to read and manage email, UNC respondents mentioned concerns regarding the management of email, including deletion and retention, more often than their peers at the private institution. Duke respondents mentioned concerns regarding security and software issues more often than their peers at the public institution.

The findings from this initial survey identified the major issues we should address during phase two of the project - the in-depth interviews. From February to June 2003, we conducted in-depth interviews with a total of 100 individuals - fifty per campus. These interviews focused on questions regarding the appraisal, authenticity, arrangement, description, physical preservation, privacy, and security of electronic files and email messages. We also interviewed approximately 25 information technology staff members in order to determine the information technology point of view regarding email and file management on each campus.

A cursory glance at the interview data shows that people at all management levels are tremendously insecure about their abilities when it comes to managing electronic files. Our goal is to arm the average user with management strategies that will both alleviate their organizational difficulties and fulfill archival and records management obligations. First and foremost, the project team wants to understand how people are interacting with their electronic files and email; then to build on those insights with the intention of empowering users through development of guidelines and instructional materials that will enable them to become more efficient and effective information managers. This research garnered a lot of data, but perhaps the most important insights involve skills and attitudes towards information management and the complexities involving digital preservation.

Email Preservation

Preserving electronic records isn't just a good idea; in many cases, it's the law. In North Carolina, all paper and electronic documents created, processed, or maintained on state time at state expense are public records, which belong to the people of North Carolina and must be made available when requested either free or at minimal cost. All documents,

papers, letters, maps, books, photographs, films, sound recordings, magnetic or other tapes, electronic data-processing records, artifacts or other documentary material, regardless of physical form or characteristics, made or received pursuant to law or ordinance in connection with the transaction of public business by any agency of North Carolina government or its subdivisions (this would include all state universities and its offices), are considered public records (North Carolina General Assembly Committee on Electronic Records, 1999). The law further goes on to say that some public records such as student academic records; personnel records, medical records, etc. are protected by specific confidentiality laws and may not be made available. Finally, public records may not be destroyed without an approved records retention and disposition schedule authorizing their destruction.

State public records' laws are complicated, and even though most users are ignorant of the law's complexities, they will be the ones held officially responsible in any legal audit. In many cases, users' behaviors are contrary to best practices, and often work against their best interests in regards to this law. The Information Technology community and archivists must work together to ensure that 1) the historical record is preserved in a meaningful and legally appropriate way; and 2) users' experience with this new software is as intuitive and natural as possible.

Electronic Records and Public Records Law

Email as Electronic Record

In order for users to effectively appraise, or make accurate decisions regarding email retention, they must first recognize the value of these documents. The concept of an email being a record is a complex one. Not only does it imply the need to select, save, store and organize digital objects commonly thought of as "temporally insignificant," or primarily casual in nature - it also implies the need to select, save, store and organize these objects *appropriately* for the historical record. Our data reflects some shortcomings in user perception. When asked whether they thought the university should preserve their "work" (this would include email as well as digital files), the respondents were almost evenly split in their opinions. Those who said "no," cited a variety of issues, including the "insignificant" nature of either their email or their job. For example, a senior Sociology faculty member said, "I don't think that I produce the kinds of documents that are legally of interest, it's not like I am discovering DNA or anything..." On a related note, many users believe that archivists have better things to do with their time and money than sort through the detritus of a typical working day. Finally, respondents had the idea that truly important documents are printed and/or stored somewhere else. Respondents were surprisingly cognizant of archival issues like the need for context and provenance, the complexities of the digital environment,

and the inability of current practice to effectively deal with such an “insurmountable amount of data.”

Of those who said their documents should be saved, there seemed to be some confusion over whether this was for historical, legal, or institutional memory purposes. Of those who were interested in historical preservation, they referred to a kind of cultural heritage of digital documents, “that would be an interesting thing to look at in 1000 years,” for example; Legal interests focused mainly on the need to store medical, student and administrative records for given lengths of time; and those citing institutional memory purposes generally want to leave behind some sort of instructions for subsequent employees.

Public Records

Whether or not respondents recognize the importance of their digital documents, the overwhelming majority (81%) recognize that there are records they must keep for legal purposes. Given the fact that 68% responded that they had never been given instructions on what kinds of documents or communications they should or should not save, they make surprisingly appropriate decisions. Users typically save: administrative and financial materials, student and patient records, official correspondence, grant materials, and professional documents like research papers and presentations. This relates nicely to those documents outlined in the North Carolina State Archives’ publication “Email as a Record in North Carolina,” which states that “a public record is any information regardless of physical form or characteristic and including electronic data processing records, made or received in connection with the transaction of public business by any agency of North Carolina government or its subdivisions.” Examples of emails that would be considered records within this context are those that contain: policies and directives, correspondence related to official business, work schedules, meeting agendas or minutes, any document that initiates, authorizes, or completes a business transaction, draft documents that are to be reviewed and/or commented upon, or reports. So intuitively, users are making some rational, dependable decisions. This is due, in large part, to training on the part of institutional archivists.

Since its inception in 1992, the UNC records management program has created individual records retention and disposition schedules for more than one-third of the University’s 1200 academic departments. The University Archives and Records Service offers workshops on a variety of topics including: records, scheduling, archiving historical records, filling, and managing electronic records.

Duke currently has no records management program, although one is in development. The state of records management at the two universities is significant for the data, because the questions regarding records retention at Duke were not asked of those employees. Because Duke is a private institution, employees do not have to comply with the North Carolina Public Records Law, or follow any specific state guidelines, whereas UNC's employees do.

Authentic Preservation

In addition to retaining appropriate records for retention, users and archivists must also take pains to preserve these objects in such a way that they do not lose their evidential value. The practice of archival preservation and organization has evolved over hundreds of years to leverage the power of context and provenance in striving for the most authentic, reliable objects. The archival tradition is in a state of flux, though, because the traditional methods of organization and description are becoming less and less relevant as the networked environment becomes more powerful and ubiquitous. Be that as it may, archives and manuscript departments must continue to collect digital objects the best way they know how.

Until there's development of more technologically innovative solutions, users and archivists will have to work within the existing frameworks. Many state archives and records programs have created "best practice" guidelines for how to manage and preserve electronic files (e.g., (State of Ohio)). In general, these guidelines cover topics such as "what constitutes a record" and "managing electronic records as one would paper files," regardless of any previously developed individual information management habits. In order to preserve records that have permanent legal and evidential value in an electronic format, accepted archival and records management practices generally refer to three categories of requirements:

1. *Hardware Formats*: Records must be stored on a reliable storage and retrieval mechanisms;
2. *Context*: Systems must maintain the context and attributes (including metadata) of the original document;
3. *Comprehensible Organization*: Emails must be organized in a system that allows others to determine the general topics to which the messages relate; and they must be clearly identifiable as archival and part of a records series

How are employees using the technological tools at their disposal - mainly their email software - to fulfill such archival guidelines?

Hardware Formats

North Carolina State Law provides general guidelines that will produce the most archivally sound documents. First, electronic records must have reliable storage and retrieval mechanisms. Arguably, there is no such thing. We neither know how to effectively or efficiently migrate or emulate electronic files for use in the future, nor do we know how much such processes will cost. The guidelines note “the practice of storing email messages with long-term value on machine readable media such as CD-ROM, 3480 tape, or digital linear tape presumes that the hardware and software required to read the data will exist in to the future.” This is not necessarily a good presumption. Moreover, the Office of Archives and History reserves the right to “accept into the State Archives stored only on those media it has the ability to read” and that it might “delegate the responsibility of long term maintenance and preservation to the creating agency.” Our data suggests that users do not intuitively recognize the instability of electronic media. When asked, “how do you store important messages,” 39% of the respondents print out the message, 17% store the message electronically somewhere else; either on a hard-drive [65%], network drive [27%], or they burn it to some other physical medium like CD or DVD [8%]. Forty-four percent save the important messages in their email program, either letting it sit in the inbox [65%] or filing it within subfolders [35%]. Of the 44% who save the messages in their email program, 30% cite flagging the message as “important” as the primary means of differentiation. Of those people who do not store important messages on their email programs [56% of the respondents], 9% accomplish this task by cutting and pasting information from the email into a Word or text document and either printing it out or saving it to their hard or network drives. Although not explicitly asked, we would assume that the rest of the people who save their messages electronically do so in proprietary email software formats.

Email Organization and Management

The second recommendation the State Archives of North Carolina makes is that email messages “must be organized in a system, so that one may determine the general topic of which the messages relate.” In other words, a large number of email messages in an undifferentiated mailbox or folder, lacking organizational and identifying characteristics will not be accepted. For this study, we divided the concept of email management and organization into to three basic components: folder naming schemas, subject line determination, and volume.

As Whittaker and Sidner (1996) observe, the organization of email messages into some sort of system can help some users employ their own personal information management strategies. While this is not a natural behavior for everyone, organizing messages consistently in meaningful folders is presently fundamental to archivists’

ability to provide intellectual access to the emails in the future to unanticipated audiences by maintaining topical and contextualized access points.

In their 1996 study of email management behavior, Whittaker and Sidner classified their subjects as “no filers” (33%), “spring cleaners” (39%) and “frequent filers” (28%). Significantly, 67% of the “no filers” were managers and many indicated this approach led to “email overload” with cluttered inboxes and lost and unanswered messages, but that high message volume, time constraints, and time spent away from their computer made filing impractical. These individuals averaged over 3,000 uncontextualized messages in their in-boxes 9 years ago.

Of the people we interviewed at both institutions, 88% are using some sort of folder system to organize their emails. Some have relatively intricate organizational schemes, with the majority of people having between 11 and 50 folders; 32% having over 50 folders and 14% having 10 or fewer folders. Thus, approximately a quarter of these academic users are either “no filers” or have very few folders, and as with the participants in the Whittaker and Sidner study, pose serious challenges to archival management of these messages.

The group that organizes their emails into folders tends to spend the most time managing their email, and tends to be the most concerned (even anxious) about their email management style. They also appear to choose appropriate types of emails to save. For example, they have folders devoted to job duties, administrative functions, personal emails, research, professional functions, teaching, and the very illuminating category: “things to keep and do.” They separate content by subject, sender, time-periods, and “function.” However, the naming of such folders is erratic, with some folders having clearly identifiable subjects (“Travel Reimbursements 2004”), and others being less intuitive, either devoted to the sender’s name (“Jane Smith”) for example, or chronological dates (“April 2004”). To the outside observer it is not always easy to determine what the content of these folders will be although the user who set up the system has little difficulty retrieving emails when they need them.

In addition to ambiguous folder names, the subject line of the email often does not reflect the content of that email. When asked, “how do you determine subject lines?” most responded that they based it on the topic or content of the email, but some said they selected it to catch the recipient’s attention. Many emails are sent in reply mode with a previously used subject line, although the topic may have changed. When asked if they changed such inappropriate subject lines, 49% said that they did not re-title such messages. This presents a barrier to future topic retrieval of these messages for both the recipient and any subsequent users of these files such as

archivists or colleagues.

In addition to having poorly organized messages and non-intuitive subject lines, users feel overwhelmed by the volume of their email. Fifty-two percent of respondents have between 1 and 100 emails in their inbox. However, 19% have over 1000 emails in their inbox alone. This number does not include emails filed in folders and sub-folders.

Those who do not use folders either have a lot of messages in their inbox or very few. Similar to Whittaker and Sidner's finding, 25% of these "no-filers" have over 3000 messages, while 25% have between 200 and 1,000 messages, and half have fewer than 20. This is an interesting finding, because those people who do not maintain an extensive folder system to organize their email also seem to be the most satisfied with their email management style. From a records management point of view, however, the lack of folders, ambiguous folder names, and inaccurate subject lines make it difficult to clearly identify which emails have archival value.

Context and Attributes

Personal information management strategies make life easier for the user. They provide the means for coping with the management of their email, but they often may not take into account the archival need to preserve the context in which an email message or electronic record was created or received. In order to preserve an authentic record, that record's attributes, such as metadata, attachments, proof of receipt, and delivery, must be kept intact.

Contextual aspects, such as value, document attributes, timelines and disposition often guide how an individual organizes their documents. From an archival perspective, these contextual characteristics provide evidential value to a document and preserve its authenticity.

When employees store their email within their email software, it is easy for them to maintain context and document attributes through the automatic preservation of metadata. Nearly all of the interviewed employees store their email within their email software program, which saves the metadata information along with the email. However, when asked "if you store a message electronically outside of your email software, does the header stay with the message," 76% said "no," 22% said "yes," and 2% said that they didn't know. It appears that users are not making a concerted effort to preserve valuable email header information when saving messages outside of their email programs. Because of this, future access to these documents will be hindered, as the messages will not be clearly identifiable in any way.

Another example of a contextual aspect is the relationship between an email message and a document attached to that message. Often an email provides information as to who created and sent the attachment, when it was received, and various other comments regarding the document. When asked, “do you save attachments outside of your email program,” a resounding majority, 89%, said, “yes,” and 11% said “no.” Of those that do save attachments outside the email program, 83% make no attempt to preserve any kind of relationship between the document and the email, while only 17% do. Reasons for either preserving or not preserving a relationship between an attachment and an email revolve around whether the respondent thinks the content of the email is important or not.

When preserving a relationship is considered important, 78% of the respondents claim that they save the email along with its attachment in the email software. Although this action preserves the context, it presents obvious problems for future access - email software having proprietary formats. Other strategies that users employ to maintain relationships between attachments and email included cutting and pasting the information they deem “important” into the attachment; printing both the email and the attachment and filing them together; or using explicit naming conventions when saving the attachment and email. In general, it is often difficult for users to maintain the context and attributes of their emails over the long-term.

Training

The in-depth interviews confirmed our suspicions that most faculty and staff at both institutions have not been “taught” how to use their email beyond how to send, open and delete messages. In fact, only 27% of the respondents said that they received any sort of email software training at all.

Our project team has developed a set of guidelines and training modules to help users make sense of the appraisal process, in the hopes that knowledge will reduce user stress and make for more appropriate decisions regarding which records to keep and which to delete (Managing the Digital University Desktop). Primarily interested in giving the user the tools and knowledge to make good decisions, we have taken care to make the documents valuable by ensuring readability by the layperson, and making the recommendations easy to incorporate into daily practice. Our basic premise was to view document management as “deletion management,” by helping the user distinguish between those few emails that they should keep, and the majority of emails that should be deleted.

Conclusions and Recommendations for an Electronic Records Management System

What does this mean for the Information Science and Archival communities? Although the behaviors identified in this project were applicable to all email and electronic files, and not just those files that would be considered archival and covered by the state guidelines, it is clear that there is a general lack of knowledge regarding appraisal and records management concepts. The implementation of a records management program and the archiving of materials will be difficult in an atmosphere where many employees don't even think that they create any documents that are important enough to the University to preserve. Perhaps the first point to address in any educational outreach should be regarding the importance of preservation - what types of records are important to the University and why.

Generally, university employees are not naturally behaving in a manner that would make the implementation of the current records management guidelines easy. In fact, some employees are making the task of records management more difficult for records managers and archivists by not organizing their files in folders and inaccurately naming files and folders. Few employees are "appraising" their records according to standard guidelines and most are using personal criteria that can vary widely from person to person. The organization and naming of email and electronic files are currently very personal decisions. It would be difficult to enforce guidelines that require employees to change their current organizational schemes.

Is there a middle ground between what employees are doing and what records managers or archivists want them to do? The Managing the Digital University Desktop has created best practice recommendations and training modules that will foster efficient, effective, and legally compliant records management behaviors, available on our website at <http://ils.unc.edu/digitaldesktop/> . In order to increase the chances of people actually adopting new practices regarding information management, we used our observations regarding their behaviors to create guidelines that are not too different from their current practices. With these guidelines the project seeks to merge current behaviors with procedures required by Public Records law and prevailing archival theories.

Acknowledgements

This project made possible by: National Historical Publications and Records Commission, National Archives & Records Administration, the University of North Carolina Chapel Hill, The UNC Chapel Hill Libraries, the School of Information and Library Science at UNC Chapel Hill, Duke University, and the Duke University Libraries.

Project Team Members: Dr. Helen R. Tibbo, SILS professor & co-principal investigator; Timothy Pyatt, Duke University Archivist & co-principal investigator; Janis Holder, UNC

Chapel Hill University Archivist; Frank Holt, UNC Chapel Hill Records Service Coordinator; Paul Conway, Director of Information Technology, Duke University Libraries; David Mitchell, University Records Manager, Duke University; Kimberly Chang, co-project manager; Megan Winget, co-project manager (2003-2005); and Ruth Monning, co-project manager (2002-2003).

References

Barreau, D., & Nardi, B. A. (1995) Finding and reminding: File organization from the desktop *sigCHI Bulletin* 27(3), 39-43

Boardman, R., & Sasse, M. A. (2004) *"Stuff goes into the computer and doesn't come out": a cross-tool study of personal information management* Vienna, Austria: ACM Press

Gould, J. D., Boies, S. J., & Lewis, C. (1991) Making usable, useful, productivity. Enhancing computer applications *Communications of the ACM* 34(1), 75-85

Managing the Digital University Desktop. (2004) *Email Frequently Asked Questions: Keeping & Deleting* Retrieved February 13, 2006, from <http://ils.unc.edu/digitaldesktop/FAQ/email/keeping/index.html>

North Carolina General Assembly Committee on Electronic Records. (1999) *North Carolina General Statutes: Public Records Law* Retrieved February 13, 2006, from <http://www.ah.dcr.state.nc.us/e-records/ncgs/ncgs132.htm>

Sheridan, T. B. (1989) Designing complex technology: Understanding it as of, by, and for people: Some dilemmas and what to do about them *Technological Forecasting and Social Change* 36(1/2), 89-97

State of Ohio. (2005) *Electronic Records Management Guidelines* Retrieved February 13, 2006, from <http://www.ohiojunction.net/erc/RMGuide/ERGuidelines.htm>

Whittaker, S., & Sidner, C. (1996) Email overload: Exploring personal information management of email *CHI96* 276-283