Chapter 18

Proposal for a Monument to Lost Data

Barry Mauer

As our society transfers its archives from print to digital media, an unintended consequence results; we lose a great amount of data. The effects of data loss can be profound; without access to vital data, our access to history may be severely diminished. Data loss threaten to undermine individual lives and major institutions.
This essay discusses the phenomenon of data loss in the age of digital media. I identify the National Archives and Records Administration (NARA) in Washington, DC as a site of data loss at a national level, and suggest mourning as a means of coping with data loss. To facilitate the mourning of data loss at a national level, I propose a monument to lost data that will be located at the NARA. This monument is to have an on-site component and an online component: a network of memorial entries created by individuals who have suffered data loss.

Our media is your memory. (Sony slogan).

> Attached to every person like a tiny galaxy will be the whole of his past — or what he takes to be the whole of his past. His attachment to it will constitute the whole of his present — or of what he takes to be the present. The neat, almost soundless instrument will contain all of each man's hope, his innocence, his garden. Then one by one, but with growing frequency, men will begin to lose their machines. (Merwin, 1969, pp. 130–131)

1 Introduction

As our society transfers its archives from print and analog media to digital media, an unintended consequence results; we lose a great deal of data. The effects of this data loss are profound; without access to our data, we lose our history, and thus our ability to function in the present is diminished. Data loss threatens to undermine not only individual lives but major institutions as well.

Writing and Digital Media

ISBN: 0-08-044863-1

Mauer, B., (2005). Proposal for a monument to lost data. In L. Van Waes, M. Leijten & C. Neuwirth (Vol. Eds.) & G. Rijlaarsdam (Series Ed.), Studies in Writing: Vol. 17. Writing and Digital Media (pp. 287–309). Oxford: Elsevier.

This essay aims to generate more discussion about the issues related to data loss. It examines the ways in which data loss have worsened in the age of digital media, and it discusses the types of data that are being lost. To provide a case study for discussion, I identify the National Archives and Records Administration (NARA) in Washington, DC as a site of data loss at a national level. Though this essay does not suggest ways of reducing data loss, it does suggest a method for coping with data loss; that method is mourning. To explain how mourning functions, the essay offers a discussion of monumentality — understood as the writing of collective memory — and also proposes a monument to lost data to be located at the NARA. In addition to this on-site monument, there will be an online component: a network of memorial entries on the Internet created by individuals who have suffered data loss. These entries will be composed of traces and "phantoms" of lost data — the phantom data being analogous to the phantom limbs experienced by amputees. The various entries of the monument will be conjoined into a structure analogous to the mycelium, the branching network of root-like fibers common to mushrooms and other fungi. The monument is designed to produce new ways of thinking about data loss and will reveal to people the collective dimensions of their personal data loss.

2 The Growing Threat of Data Loss

Numerous critics blame electronic media for overtaking and undermining the public sphere, that space for the free exchange of ideas which provides sustenance for a democratic society (see the work of Noble, Postman, & Mander). These critics decry the commodification of electronic communications, the invasion of private electronic media enable, and the disinformation electronic media spread routinely (see Rheingold, 1993). Less well known is another critique; electronic communications technologies hasten the destruction of data, which in turn impedes citizens from making informed decisions. This argument is counterintuitive; does not electronic media store and transmit more data than ever? Yet the flood of data in electronic societies may be the primary cause of data loss within them.

Substantial publicity about the growing quantity of stored data may be obscuring the problem of data loss. For example, a report by researchers Lyman and Varian (2003) points to dramatic increases in data and data-storage capacity:

> Print, film, magnetic, and optical storage media produced about 5 exabytes of new information in 2002. Ninety-two percent of the new information was stored on magnetic media, mostly in hard disks. How big is five exabytes? If digitized with full formatting, the seventeen million books in the Library of Congress contain about 136 terabytes of information; five exabytes of information is equivalent in size to the information contained in 37,000 new libraries the size of the Library of Congress book collections (Lyman & Varian, 2003).[1]

[1] Also see New. His study finds explosive growth of world's information is only beginning (2000). "It's taken 300,000 years for humans to accumulate 12 exabytes of information. It will take just 2.5 more years to create the next 12 exabytes, according to a new study produced by a team of faculty and students at the School of Information Management and Systems at the University of California at Berkeley."

Less public attention is paid to the phenomenon of data loss. It seems to be exceptional; disasters, such as the massive data loss that resulted from the destruction of the World Trade Center, appear to be anomalies rather than regular occurrences (Joyce, 2001). Because the quantity of stored data and data-storage capacity is growing rapidly, the public might assume that data loss is becoming a thing of the past. Not so. Major institutions have been grappling for some time with the problems posed by data loss. A 1999 study conservatively estimated corporate data loss had cost the U.S. economy $11.8 billion in 1998, and the study suggested these costs were likely to increase (Lost data cost U.S. economy $11.8 billion in 1998 and 1999).

Though we have some cause to celebrate the enormous data-storage capacity of digital media, our society has not come to terms with the enormous data loss that accompanies our shift from textual and analog archives to digital archives. This project aims to draw our attention to data loss as a profound problem affecting modern life.

We should distinguish the loss of memes — which are the primary elements of collective memory — and the loss of personal data. Loss of mathematics would be catastrophic for the society, but the loss of your checkbook, while bad for you, is unlikely to affect the society as a whole. Personal data is in greater danger of loss because it exists in fewer copies. Nonetheless, personal losses can affect individual people profoundly and can affect groups when large numbers of people lose their personal data.

All societies must pass vital knowledge to succeeding generations; therefore all societies face the vexing problem of information storage. Living memory — the memory stored in the human mind — is extremely limited; it is often unreliable and cannot survive beyond the grave; thus human societies have designed memory machines — data-storage systems — and institutions, such as pedagogy, charged with the preservation and dissemination of data. Gregory Ulmer reminds us that "Writing as a technology is a memory machine, with each apparatus finding different ways to collect, store, and receive information outside of any one individual mind (in rituals, habits, libraries, or databases)" (Ulmer, 1994, p. 16).

Units of data essential to an apparatus — any institution charged with the preservation and dissemination of specialized knowledge — are called memes. Ancient rhetoricians recognized the importance of memes (though they did not use that word) by making the first and most important of the five parts of rhetoric the inventio — an imaginary storage place for "true things" (Cicero). The inventio was a kind of archive for commonplaces, the accepted units of wisdom possessed by the society.

According to Csikszentmihalyi (1996), memes are bits of information that a society passes on to the next generation in order to replicate itself. All social institutions — such as family, entertainment, and the sciences — have specialized memes that allow them to survive and maintain continuity from one generation to the next. During the early development of a social institution, new bits of knowledge appear at a rapid rate. Those deemed essential get taught to future generations, while those deemed non-essential become dormant. The selection process governing novelty in our information world is similar to the selection processes governing novelty in the natural world.

The analogy to genes in the evolution of culture are memes, or units of information that we must learn if information is to continue. Language, numbers, theories, songs, recipes, laws, and values are all memes that we pass on to our children so that they will be remembered. It is these memes that a creative person changes, and if enough of the right people see the change as an improvement, it will become part of the culture (Csikszentmihalyi, 1996, p. 7).

Memes must be learned or they will not survive. Computers, unlike humans, can replicate and store memes with little difficulty, but they do not understand these memes. Until they do, human memory and its institutional ally, pedagogy, will bear primary responsibility for storing and transmitting memes.

How fragile are memes? A historical view reveals that the fragility of memes changes over time. The memes stored during antiquity were fragile in some ways and durable in others. The libraries at Alexandria in ancient Egypt — there were actually three libraries in the system — had a mandate to collect all the knowledge in the world. At its peak, these libraries had roughly 500,000 scrolls before they were destroyed. Texts at that time were rare items; they had to be copied by hand, a very laborious process, and many existed as single copies. When the libraries were destroyed, so went the vast majority of knowledge stored until that point. Yet the memes of antiquity were durable; the fragments of ancient culture that survived have had an enduring impact on Western culture for 2500 years. "While modern scholars often lament the amount of information lost through the centuries since the Museum's fall, an amazing number of Alexandrian discoveries and theories, especially in mathematics and geometry, still provide the groundwork for modern research in these fields" (Brundige, 1995). The ancients produced a holistic way of storing essential data; small fragments (individual texts) reflected the greater whole (the society's knowledge). Much of what survived from antiquity is intertextual and thus we can find references in surviving texts to missing texts and we can make inferences about those missing texts. Additionally, we can infer from other ancient records, such as property records, gravestones, and inscriptions on buildings, about the principles governing ancient societies.

While the memes of antiquity were fragile because of the rarity of texts, the memes of modernity are fragile because of the abundance of texts in our archives. There are two defining characteristics of the modern data loss crisis. First, there are far more memes — or units

Figure 1: The National Archives and Records Administration (NARA).

of essential data — to manage because we have more institutions, and more complex institutions that require additional memes to sustain themselves. Also, because of the volume and complexity of our stored data, essential data has become increasingly harder to identify from within a larger ocean of data that may be largely non-essential. The genre of detective fiction, which arose in the 1800s as data archives grew at explosive rates, points to problems arising from our data explosion; the detective struggles to grasp the significance of seemingly insignificant data — clues — within an undifferentiated mass of facts.

The NARA in Washington, DC is emblematic of the modern data crisis. According to Stille, author of The Future of the Past, this federal agency, which is charged with storing and making available government records, "may have lost more information in the information age than ever before."

The NARA (Figure 1) was created during the 1930s on the optimistic premise that the government could keep all of its most vital records indefinitely, acting as our nation's collective memory. Now, as it drowns in data and chokes on paper, the agency is facing the stark realization that it may not be able to preserve what it already has, let alone keep up with the seemingly limitless flow of information coming its way (Stille, 2002, p. 303).

Among the massive problems facing the NARA, according to Stille (2002), are the following:

1. A huge backlog of data waiting to be transferred from decaying obsolete media to new media.
2. The need to replace "archival" digital media every few years.
3. "Decisions about what to keep ... made by default" as large portions of the archives deteriorate.
4. The speed at which "each new generation of equipment supplants the last" is leading to early obsolescence, raising the need to transfer databases, yet again, to a newer media.
5. "The newer the technology, the greater its fragility."
6. Competition for budget resources leading to tough choices between storage and dissemination.
7. Lack of physical space.
8. An increase in the amount of paper as government agencies print their computer files.
9. An increase in the amount of electronic data to be archived.
10. "Different kinds of computers, software programs, and formats."
11. Many files, such as email, existing in formats "not designed with long-term storage in mind."
12. The downsizing of federal agencies leading to the massive transfer of those agencies' records to the NARA.
13. Budget cuts at the NARA.
14. Refusal on the part of government agencies to be selective about what they add to the archive.

Stille cites a 1996 study of the NARA concluding "that, at current staff levels, it would take approximately 120 years to transfer the backlog of nontextual material (photographs, videos, film, audiotapes, and microfilm) to a more stable format" (Stille, 2002, p. 300). The 1996 study examined only non-textual materials already in the archives' collection; it does not account for new materials to be added to the collection. The collection of textual materials at

the NARA is growing more rapidly as a result of a 1989 case in which "a public interest group trying to get information about the Iran-contra scandal successfully sued the White House to prevent it from destroying any electronic records. The result is that all federal agencies must now preserve all their computer files and electronic mail" (Stille, 2002, p. 304). However, Stille points out, "Because government offices use different kinds of computer files, software programs, and formats, just recovering this material has proved to be a logistical nightmare" (Stille, 2002, p. 304). If the NARA has difficulty preserving the documents from one of the most famous scandals in U.S. history, how can we be certain that other kinds of data, which may be vital to our future well being, has been preserved?

Stille concludes his chapter about the NARA with the following:

> There is not likely to be a modern Sophocles in the databases of the Department of Agriculture or the Census Bureau. The greater risk, instead, is of such a vast accumulation of records that the job of distinguishing the essential from the ephemeral becomes more and more difficult. The Archives of the future may resemble the "Library of Babel" that Jorge Luis Borges imagined nearly sixty years ago, an infinite library that contained every conceivable book in the universe. There were books that consisted purely of a repetition of a single letter of the alphabet and others in which all the pages but one were blank. The discovery of an intelligible sentence was cause for jubilation. Eventually, after many centuries, the librarians of Babel were driven to despair in their unfulfilled quest for a coherent, complete book. (Stille, 2002, p. 309)

Borges' "Library of Babel" suggests that archived data is effectively "lost" if we are unable to access it. There are enormous problems related to making searchable the data in electronic archives, not least of which is the problem of "tagging" (with SGML, XML, and other forms of searchable coding) the images stored in electronic databases. Suppose a researcher were to try generating a list of movies that contain images of apple blossoms; should the researcher assume that someone has tagged all such images in a searchable database? It would be impossible to generate all possible search tags for even a relatively limited number of images.

The Internet also draws criticism as a vast wasteland of disorganized and unsearchable data:

> The Internet has been described as a library where at the moment there is no catalogue, books on the shelves keep moving, and an extra truckload of books is dumped in the entrance hall every hour. Unless it is properly structured and constantly monitored, the positive feature of radical decentralization of knowledge will degenerate into a medieval fragmentation of the body of knowledge, which in turn means a virtual loss of information. Already it is no longer possible to rely on the speed of our networked tools to browse the whole space of knowledge and collect our information in a reasonably short time. If global plans are disregarded or postponed and financial commitments delayed, the risk is that information may well become no easier to find on the network than the proverbial needle in a haystack. (Floridi, 1995)

Though Floridi might hope for a global effort to structure and monitor the Internet, that dream seems increasingly unrealizable. Optimists will argue that science and commerce will eventually solve the problems of data management by building bigger and better hard drives, improving software for archiving and searching, and training users to back up their data. Pessimists will argue that even the best science will not save the world's disappearing languages, that advances in technology so far have exacerbated some problems of data loss, and that there will always be human error as long as there are humans.

Are modern societies in danger of losing memes? Certainly more data is stored today than has ever been stored. Oceans of data fill our libraries and computers. In general, we no longer face the problem of data "scarcity" that the ancients faced. A television show or a photograph or a written text can exist on a million computers and discs. Some of the essential texts of our culture, such as the "Declaration of Independence," "Huckleberry Finn," and "Star Wars," exist in so many copies and in so many forms that there is little danger of a catastrophic loss that would wipe them all out. There is sufficient backup to ensure the survival of this data for generations to come. Yet even this familiar data may be corrupted: the original parchment document of the "Declaration of Independence" is faded to the point of indecipherability; the manuscript for "Huckleberry Finn" was lost for over a 100 years while popular versions with hundreds of variances (errors?) from the manuscript were widely circulated; and George Lucas has released "Star Wars" after making many changes to the images and the editing. People will debate the significance of these changes to major texts in our culture for years to come.

Even though some memes are widely available, that does not guarantee their survival. Only if people understand how to access and use them will memes survive. The knowledge required to access and use any data is itself composed of memes that could disappear under certain conditions. For instance, the way we search databases has changed dramatically, and with potentially dire consequences; "For the last few years, librarians have increasingly seen people use online search sites not to supplement research libraries but to replace them" (Hafner, 2004).

> "Google searches an index at the first layers of any Web site it goes to, and as you delve beneath the surface, it starts to miss stuff," said Mr. Duguid, co-author of "The Social Life of Information." "When you go deeper, the number of pages just becomes absolutely mind-boggling." (Hafner, 2004)

In some ways, we have overcome the problem of information scarcity, but the ability to navigate and interpret the oceans of data we store poses as a significant problem for us as information scarcity once posed to earlier generations. Our colleges and universities struggle to train the next generation of researchers, in part because research methods keep changing so quickly.

Sometimes the loss of personal data does have major consequences for collectives. During the Serbian ethnic cleansing campaign in Kosovo, the Serb militias stripped the Kosovars of all identifying papers and photographs so that they could not prove who they were and where they had lived (Rueb, 2004). There was a systematic effort to deprive Kosovars of their history, their identity, and their claim to their property (see Figure 2). As such, the individual losses, taken together, were a catastrophe for the entire society, not just for each Kosovar.

Figure 2: A British KFOR officer sifts through a pile of Kosovar Albanian passports and ID cards, which had been confiscated by Serb forces and were found on June 13. The lack of ID papers is complicating the OSCE's task of registering voters (Reuters photo. Image and text from Rueb, 1999).

Increasingly, data loss threaten institutions as hackers and viruses cripple government and corporate sites and block public access to essential data. Hard drive data losses also result from power outages, storm damage, fire or explosion, hardware/software error, flood and water damage, earthquake, network outage, and human error.

A data protection industry, composed of scientists, engineers, trainers, and archivists, has arisen to address these problems. The data protection industry tries to reduce data loss through improved training, security, archival systems, and recovery techniques. Despite refinements in these techniques, the problem of data loss remains and seems unlikely to diminish. "Eventually every hard disk or backup will fail" (DataMechanix, 2004).

The data recovery industry[2] is able to rescue data from most of the crashed hard drives they receive (see Figure 3). But professional data recovery is relatively expensive — typically the cost of data recovery greatly exceeds the cost of the hardware — and the data

[2] See Lost in space: Data recovery (2001).

Figure 3: Life or death: a hard drive surgery (http://www.datamechanix.com/dm_clnrm.jpg).

recovery process can be extremely anxiety producing for clients. At least one data recovery company has hired counselors to help its clients deal with data-loss anxiety (see Mieszkowski, 1997).

The project described here — the monument to lost data — is relevant to those cases in which data cannot be recovered and must be considered lost. In these cases, it is appropriate and healthy to embrace a strategy for coping with loss: mourning. Mourning is the process whereby one achieves a measure of detachment from a lost person or object. (Ulmer, 1994). Mourning follows grief, which is the immediate reaction people have to a profound loss such as the loss of a loved one, but also social and symbolic losses (divorce, loss of a job, etc.).

The process of mourning, in addition to its healing effects, creates identity in powerful ways. Children learn to adopt the values of their parents and grandparents through the introjection of their elders' ego ideals. When elders die, children learn to cope with the loss by retaining the values of those who are gone. Mourning functions as a species of memory, creating continuity out of contingency; it puts loss into perspective and allows the young to adopt the values of the dying generation.

The process of introjection, however, is usually uncritical; most of us do not choose our ego ideals. But mourning can have a critical component. The monument to lost data foregrounds critical reflection in the mourning process; I explain this critical component, which includes reflections about our choices of values, in greater detail at the end of this chapter.

I support the mourning of data loss as a way of dealing with the special conditions of our information age. The most effective way to deal with data loss, I believe, is to recognize it as a collective experience and not just a personal one. A collective approach to mourning lost data provides greater perspective on the problem and will enable social networks to function as support groups for individuals who have suffered data loss.

In the previous chapter of this volume, Selfe and Hawisher examine the wide variations in digital literacy and access to digital technologies throughout the world. One result of political, economic, and technical inequities, is that collective entities — cultures, nations, and institutions — are losing essential data; some of the world's languages, dialects, rituals, oral histories, neighborhoods, and buildings are disappearing. Although components of this data can be recorded by ethnographers and historians, it is in danger of being lost as "lived memory" and thus as memory essential to a culture. According to linguist Michael Krauss, approximately 3000 languages will become extinct in the 21st century (Shorris, 2000). Reservoirs of traditional knowledge — the "data" stored in dances, songs, rituals, and myths — are disappearing from the planet, in part because the cultures and institutions in which they reside are not integrated into the global information network. Krauss implies that these losses cannot be calculated fully: "Each language is a unique repository of facts and knowledge about the world that we can ill afford to lose, or, at the least, facts and knowledge about some history and people that have their place in the understanding of mankind. Every language is a treasury of human experience" (Kolbert, 2005). When traditional cultures transfer their wisdom to electronic media, as many do, they trade one set of problems for another as they lose ritual practices that provided continuity over generations; they also face the same problems of data loss that other electronic cultures face.

Not only are our data archives unstable; the terms we use to discuss data are unstable as well. The definition of "data" itself has changed considerably in 30 years, becoming associated less with humans and more with machines. The 1971 edition of the OED defines datum, the singular of data as, "A thing given or granted; something known or assumed as fact and made the basis of reasoning or calculation; an assumption or premise from which inferences are drawn" (Datum, 1971, p. 648). This 1971 definition refers to data in human terms. Later, by contrast, connotations of automatism take over the definition. A recent entry, found in WorldCom's Communications Library, defines data as, "Any coded information that can be processed or output by a computer or other machine" (*Data*, 2001). Only machines use data, according to WorldCom, not humans — unless WorldCom was slyly admitting that we have become post-human. In this definition, machines process or output data; they neither "know" it, nor "assume" it, nor do they draw inferences from it. Ironically, since the collapse of WorldCom into bankruptcy, the website with WorldCom's definition of data is no longer available. Prior to WorldCom's definition, a government website had defined data as having human and non-human attributes; "Representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means. Any representations such as characters or analog quantities to which meaning is or might be assigned" (*Data*, 1996). My use of the term data is intended to be as broad as possible, encompassing both the human and non-human connotations raised by the term. Thus data loss, for my purposes, means the loss of a hard drive, the loss of a spoken language, and the loss from Alzheimer's of a person's living memory.

3 Artists Address the Shift in Data-Storage Practices

Many of the emerging issues related to data storage and data loss have been addressed in the arts. Andy Warhol and Merwin, both active in the 1960s, were among the first people

to grapple with the implications of our changing data-management practices, specifically the practices involved in "passive diary keeping" (Johnson, 2005).

Warhol recorded all his encounters, no matter how mundane, using photographs, audiotape, videotape, and film.

> The year 1969 began with a flurry of ideas. What about a television show, Andy suggested, called "Nothing Special," consisting of six hours of people walking past a hidden camera? (He had recently installed one in the Factory to improve security.) ... At Brigid Polk's suggestion, he had started taking Polaroids of penises — of any Factory visitor he could persuade to drop his trousers. He later estimated that he had taken thousands of such pictures ... Indeed, Andy's desire to record everything around him had become a mania. As John Perrault, the art critic, wrote in a profile of Warhol in Vogue: "His portable tape recorder, housed in a black briefcase, is his latest self-protection device. The microphone is pointed at anyone who approaches, turning the situation into a theater work. He records hours of tape every day but just files the reels away and never listens to them." (Bockris, 1989, pp. 246–247)

Also in 1969, Merwin, in "The Remembering Machines of Tomorrow," foretold a future of ubiquitous memory machines that would feature limitless data-storage capacity.

> Attached to every person like a tiny galaxy will be the whole of his past — or what he takes to be the whole of his past. His attachment to it will constitute the whole of his present — or of what he takes to be the present. The neat, almost soundless instrument will contain all of each man's hope, his innocence, his garden. Then one by one, but with growing frequency, men will begin to lose their machines. (Merwin, 1994, pp. 130–131)

Merwin offered two critiques of the memory machines: that the machine's archives of the past and present are composed of misperceptions, and that the loss of data stored in the machines will be crippling. Merwin painted a picture of the future that resembles our present: ubiquitous media, limitless information storage, and total surveillance. His model for the person of this future may have been Andy Warhol, ca. 1969, with one key difference; Warhol never accessed the data he had recorded. He filed it away and no one accessed his collection until after his death.

Warhol's media collection includes more than 4000 audiotapes, consisting for the most part of everyday conversations. Warhol made himself the hub of a recording apparatus, though the recordings he produced could have been made, in theory, by anybody; Warhol was part of a machine, a "factory," that produced citation. What Warhol recorded was nothing in particular and what he collected was never accessed. Warhol refused to distinguish between essential and ephemeral data. Merwin's futuristic people, however, selectively record their memories, creating inaccurate impressions that substitute for actual experience. Their memories result from a type of wish fulfilment.

Passive diary keeping, as imagined by Warhol and Merwin, is now a part of daily life for ordinary people.

> The cell phone manufacturer Nokia recently introduced a new software package for camera phones and Windows PCs called Lifeblog, which combines e-mail and the passive diary mode of the photoblog in one artful package. In essence, Lifeblog records a timeline of all the events that flow through your cell phone's memory. Schedule an appointment, and Lifeblog will put it on the timeline; take a picture, and Lifeblog will archive it; get an instant message from a friend, send an e-mail, or retrieve a voice-mail message — Lifeblog will store it away in its running account of your digital life. When you sync your phone with your PC, you can launch the Lifeblog program and see a rendered account of your time — a long thread of information, woven together with images you've captured along the way. The premise behind Lifeblog is not a new one: A number of computer visionaries, including Gordon Bell of Microsoft, have proposed software interfaces organized fundamentally around time, as a way of augmenting memory. Bell's experimental project, called MyLifeBits, chronicles the entire flow of information through his life — everything from articles and books to phone calls and home movies. That kind of data storage might seem overly ambitious if you prefer to spend your time living your life and not archiving it. But the beauty of Lifeblog's orientation around the cell phone and the camera is that it lets you record life away from the PC screen, without actively thinking about the archival process. (Johnson, 2005)

These new technologies make the worlds of Warhol and Merwin real for ordinary people. But Warhol, and Merwin's futuristic people, pose problems for us; if we choose to store everything without making distinctions, as Warhol did, the very quantity of our stored data renders it unusable. If we selectively record data, we may, to our peril, miss something vital. If, like Warhol, we never access our data, then we may not need to worry much about the consequences of losing it. If, like Merwin's futuristic people, we become utterly dependent on our memory machines in order to function, then we should worry very much about losing them.

4 The Genesis of an Idea: Mourning Data Loss

For several years before I conceived of this project, I had been interested in mourning and its relationship to electronic media (Mauer, 1996). I had intended to construct a monument that would demonstrate that the electronic realm could support mourning, but I had not identified a problem — a sacrifice to memorialize — worth the effort. Then one day, when I walked into the English Department at the University of Central Florida, I found a worthy problem; the department secretary's computer had crashed. The hard drive was lost and with it was the institutional history of the department. Many department files existed in

backup copies on paper, but some data had not been backed up. The archival structure of the data, the nested folders that stored the documents, had disappeared along with the data.

The secretary was distraught; she did not know how profoundly the loss of her hard drive would affect her and would affect the department. When we lose data, I realized, we suddenly confront the fact that we are dependent on having access to it and may be crippled without it. I asked the secretary whether she would mourn the loss of her data. The idea intrigued her and we set up a time to discuss it. During the discussion, she made the following remarks:

1. Her natural memory was useless for replacing the data on the computer. There was no way she could recall the contents of all of the documents that had been stored on it and it would be impossible for her to reproduce the documents from natural memory.
2. Before her computer crashed, she had known of others who had lost data on their computers. Still she had neglected to back up her work.
3. She retained the archival structure of the data in her natural memory. If called upon to retrieve a document, she could remember the structure of nested folders and the pathways she took through the structure in order to retrieve the document. Her picture of the archive remained intact even though the archive itself no longer existed.

The secretary, as the person in charge of storing and retrieving the department's data, had a relationship to this data that was both personal and institutional. I wondered whether there was a deeper connection between the two types of data loss — personal and institutional — and whether the loss of personal data was part of a larger pattern, something with a collective dimension. I also wondered whether the experience of data loss in the modern world differs significantly from the experiences of data loss that had come before. These topics are too broad to be addressed adequately in this single chapter (a book by the author is planned). The monument I propose below seeks to create new connections between personal and collective forms of data loss.

5 Monumentality

Monumentalists — those who seek to guide collective mourning through the production of monuments — face a critical process. They must ask themselves which losses ought to be considered as sacrifices. They must ask themselves which ideals these sacrifices honor. The design of a monument reflects the monumentalist's consideration of these questions. The proposal for this monument recognizes lost data not as an accident, an avoidable mistake, but as an unavoidable loss that we may choose, if we so desire, to designate as a sacrifice. My hope is that we will recognize data loss as a sacrifice and that we will see this sacrifice as a price we pay for our collective values and behaviors. We might then choose to reconsider the wisdom of our collective values and behaviors in relation to the storage of memes. The monument to lost data will have three main functions:

1. To explore the relationship among our values, our behaviors, and our losses pertaining to data.
2. To reveal the links between personal data losses and collective identity.
3. To forge new identities around shared data losses.

The monument will support the work of mourning lost data, specifically the data loss at the NARA, but also the data loss experienced by ordinary people in their daily lives. The monument will allow for a wide variety of personal relationships to the monument because it will be constructed mostly from personal contributions, much like the Names Project — or "AIDS Quilt" — which is a gigantic monument made entirely from individual squares made and contributed by those mourning the loss of a loved one to AIDS.

Monumentality does not seek ways to avoid loss, though it has no quarrel with rationalist efforts to reduce or eliminate data loss. Monumentality aims to represent the values for which the losses occurred. Values are determined by the price we are willing to pay to sustain our behaviors. What values might be honored by data loss? A provisional answer: thousands of people suffer data loss because our society demands progress, which it defines as increased efficiency and storage capacity. Efficiency and capacity are values we are willing to pay for and pay for dearly.

Can data loss be considered sacrifice if its victims do not voluntarily destroy their data? Yes. Human beings voluntarily invest their labors in data-storage systems that produce data loss "victims"; thus we should acknowledge our willingness to sacrifice. To better understand this formulation, let us consider the loss of soldiers to war; soldiers do not volunteer to die on the battlefield, yet we recognize battlefield deaths as sacrifices because the battlefield is an environment in which death is more likely and our society willingly sends soldiers to battlefields to die. Auto deaths are a related example; people do not voluntarily die in auto accidents, yet in the United States we willingly accept the inevitable deaths of roughly 35,000 people per year for our right to own cars and drive them whenever we like.

Any monument to lost data will be "abject," meaning that data losses are not, as yet, part of the recognized category of national loss; they do not yet constitute our sense of national identity as do the deaths of soldiers during war. Because an abject loss is degraded and debased, the work of mourning abject losses is more difficult; there is little acceptance of it and support for it. Gregory Ulmer, who invented the concept of abject monuments in his essay "Abject Monumentality," proposes two abject monuments, one for auto deaths and one for pet deaths. His monuments called for recognition of sacrifices on behalf of our right to drive cars and own pets. Unlike traditional monuments, which mark the sites of the officially sacred, abject monuments mark the sites of sacrifices not yet accepted as sacred; in particular, abject monuments, according to Ulmer, mark the sites of degraded losses, such as auto deaths. How do we know auto deaths are degraded? Ulmer refers to a piece in The National Review criticizing the design of the Vietnam Veteran's Memorial Wall, claiming that the list of dead soldier's names degraded their sacrifice for the nation: "THEY MIGHT AS WELL HAVE BEEN TRAFFIC ACCIDENTS" (Ulmer, 1993, p. 8).

An abject monument claims a degraded loss as a sacred expenditure made on behalf of the collective. For example, auto deaths are necessary for the nation because those who died in auto accidents were exercising the right to drive their cars anytime, a right for which the nation is willing to sacrifice thousands. The proposed monument to lost data, also an abject monument, will recognize data loss as a sacrifice made on behalf of the nation's desire for progress, understood as our desire to augment our memories with machines.

Bataille (1988–1991) described sacrifice as "the antithesis of production, which is accomplished with a view to the future; it is consumption that is concerned only with the

moment" (p. 49). Through sacrifice, according to Bataille, we transform the object from the realm of the profane — characterized by "use value" — to the realm of the sacred — characterized by "uselessness." At once destructive and creative, sacrifice unites the community, usually through ritual, by removing its objects and its members from the world of profane things and bringing them into collective, intimate contact with the sacred.

Modern people typically provide no place for sacrifice in their daily lives. As a result, Bataille argues, they do not acknowledge their expenditures as anything other than losses within the restricted economy of wealth accumulation. By neglecting sacrifice, modern people suffer a loss of intimacy and community. The goal of making a monument to lost data is to turn the destructive character of data loss into a creative one by uniting the community in its sacrifice and helping it to understand and maybe re-evaluate its core values.

In "Abject Monumentality," Ulmer (1993) argues that the electronic sphere needs its own practice of monumentality to supplement the monumental practices of literature and architecture. He proposes electronic monumentality as an answer to the claims "made by modernists" (especially in architecture) that electronic technology was largely responsible for the decline of the public field.

> The monumental function of architecture included mourning, understood as the collective version of the psychology of identification — the formation of the superego in an individual through the internalization (introjection) of ego — ideals. Monumentality was responsible for maintaining a sense of national identity from one generation to the next (hence the mourning by one generation for the loss of the previous generation, back to the Founding Fathers). (Ulmer, 1993)

Monuments, by Ulmer's definition, are places and objects that remind people of sacrifices they make to maintain their behaviors and values. They have the following features:

1. Monuments designate a set of deaths or losses as significant and ascribe to them a great value. A classic example: the historians of the French Revolution, Michelet and Renan, represented the dead of the anti-Huguenot pogrom of 1572 as fratricidal sacrifices necessary to the development of the French nation, though at the time of the pogrom the "nation" would not exist for hundreds of years, and the killers and their victims would never have understood themselves as "Frenchmen" or "brothers" belonging to the same nation. Yet these ancient slaughters became "family history" for the French historians of the 18th century. Michelet and Renan demonstrate that monumentalists put words in the mouths of the dead, claiming to reveal their "true" desires. Victims of wars become sacrifices for "the Nation ... even when these sacrifices were not understood as such by the victims" (Anderson, 1991, p. 41).
2. Monuments identify specific deaths or losses within a large narrative framework as sacrifices on behalf of shared collective values. For example, during wartime, nations honor soldiers who have died so that the nation might live. Monuments ensure their eternal life in citizens' memories, even if it is only as an anonymous victim represented by a Tomb of the Unknown Soldier.

3. The living members of a society become indebted to those who have sacrificed themselves on their behalf. At the least, that debt is to mourn them and to retain their memory. Beyond that, members of the next generation are obligated to sacrifice themselves to honor those who sacrificed themselves in the past.
4. Mourning and sacrifice enable societies to transform the contingencies of life into continuity and to transmit preferred social values to the next generation.
5. Because monuments shape memories, values, and behaviors for succeeding generations, people can take more active roles in social transformation by learning how to create monuments of their own.

In his "Abject Monumentality" essay, Ulmer (1993) does not make explicit the poetics of abject monuments, other than to note that they should be attached as "asterisks" to existing monuments. Thus the design features I offer below are intended to be specific to this monument, though they may be applicable to other monuments.

6 Phantom Data

It may come across as paradoxical and even perverse to propose a monument to lost data since what should we put in it but data? However, the data stored in the monument will be of special types — it will be the traces and the phantoms of lost data. Traces include fragments of lost texts, while phantoms include recollections of texts for which no traces survive. Traces of data might be broken hard drives whose data had been destroyed, texts in archaic languages, and found photographs. Found photographs are mysterious because they have been detached from their original context and there may be no "Rosetta Stone" — which provided researchers with translations of ancient Egyptian text into Greek — that enables us to understand their significance. For example, people may possess ancient family photographs yet have no knowledge about the people in them.

Phantom data differ from trace data. In the discussion I had with the secretary whose hard drive crashed, she explained that she had retained the archival structure of the hard drive data in her natural memory. When called upon to retrieve a document, she could remember the structure of nested folders and the pathways she took through that structure in order to retrieve the document. Her mental picture of the archive remained intact even though the archive itself no longer existed; she was haunted by the data she had lost and this haunting caused her emotional pain. I coined a neologism to describe the secretary's condition — "phantom data pain" — which can be explained by an analogy to phantom limb pain.

> "Phantom limb" pain has been recorded almost as long as people have been losing limbs and surviving. As we can see on the motor cortex, specific areas function to map out specific parts of the body. Losing a part of the body doesn't necessarily stop the cortex from continuing to "map" the missing part, adding a slight twist to Korzybski's, "the map is not the territory" (Austin, s.d.).

A common treatment for phantom limb pain is the Ramachandran Method. The practitioner uses a mirror and a box to treat the patient (see Figure 4). With the healthy arm through the hole corresponding with the reflective side of the mirror, he maneuvers his phantom arm through the other hole (be imaginative with this one to help the client do this). Thus now, the client has a reflection of his real arm that, with some maneuvering, will correspond exactly with the missing limbs' location. This enables the brain of the person to achieve feedback to the motor area of the brain corresponding with the phantom:

> Philip rotated his body, shifting his shoulder, to "insert" his lifeless phantom into the box. Then he put his right hand on the other side of the mirror and attempted to make synchronous movements. As he gazed into the mirror, he gasped and then cried out, "Oh, my God! Oh, my God, doctor! This is unbelievable. It's mind boggling!" He was jumping up and down like a kid. "My left arm is plugged in again. It's as if I'm in the past. All these memories from so many years ago are flooding back into my mind. I can move my arm again. I can feel my elbow moving, my wrist moving. It's all moving again" (Austin, s.d.).

Ramachandran asked Philip to close his eyes and Philip's phantom arm once again became lifeless until he once again opened his eyes.

Figure 4. The Ramachandran Mirror Box. Austin, A.

Curiously, despite the historical referential experiences that flooded into consciousness, 4 weeks later following 10 min a day with the box and mirror, Philip reported that the limb had gone, "all I have now is my phantom fingers and palm dangling from my shoulder." The pains had significantly reduced (only the fingers still hurt — the rest had gone) and Philip now possessed an altered but more realistic body "image" mapped onto his sensory cortex:

> "It's not clear why his fingers didn't disappear, but one reason might be that they are over-represented — like the huge lips on the Penfield map — in the somatosensory cortex and may be more difficult to deny" (Austin, s.d.) (see Figure 5).

Ramachandran's research into phantom limbs provides the following insight; although a limb may be missing, it is still represented in the mind. In both cases, phantom limbs and phantom data, we continue to receive feedback about something that is no longer there. The monument to lost data is a variation on the Ramachandran Method, designed to help us put the loss of our data in perspective and to achieve a more realistic picture of our situation in relation to augmented memory. In the Ramachandran Method, the patient uses the reflected image of a healthy limb to gain mastery over the feedback that the mind senses from the missing limb; the amputee learns to accept a new body image, and this new cognitive map of the body reduces suffering and increases adaptability.

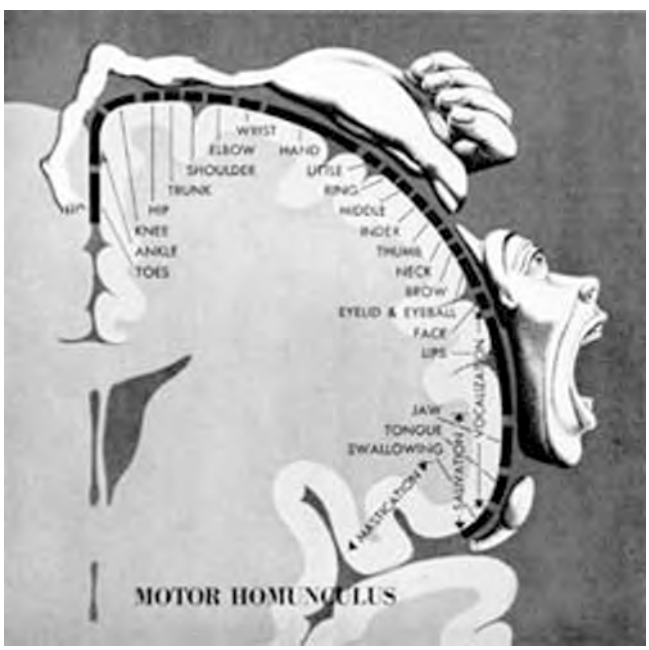

Figure 5: The Penfield map (also known as the homunculus) (http://www.eatonhand.com/hw/homunc.gif).

To help the mourner gain mastery over lost data, the monument to lost data uses a process analogous to the Ramachandran Method. In this variation, the mourner recollects and reflects upon the lost data by reference to stored data, in other words, to data that has not been lost. For example, a missing family photo album may be mourned by reflecting upon a different family photo album, perhaps an album from a different family. The loss of a hard drive may be mourned by reflecting upon the contents of another hard drive. In each case, the mourner looks for patterns, highlighting details in the stored data that resemble details in the lost data. In this way, the mourner regains some control of the feedback that the mind senses from the missing data.

7 Designing the Monument to Lost Data

Although I have put together some of my ideas, outlined below, about the location and design of a monument to lost data, there should be community deliberation over these issues. Such deliberation would raise debate about the values embodied by various options, as well as the community's investments in the monument itself. My design ideas are meant as contributions to a discussion that I hope will take place at a national level.

In my proposed design of the monument to lost data, there would be two components: one physical and one virtual. The physical part would consist of a corporeal monument at a public location, while the virtual part would consist of an electronic database accessible anywhere in the world by anyone with an Internet connection. I propose a corporeal monument in the shape of mushrooms to be located at the NARA, along with a worldwide virtual monument, accessible via the Internet, to be modeled upon the mycelium (see Figures 6 and 7). The mycelium, the branching organ of mushrooms and other fungi, is a rhizomatic structure of networked passages that gives rise to new species of fungus — in this case, new forms of memory for diverse data losses.

I chose the mushroom/mycelium as the structuring metaphor for the monument because "Fungi play vital roles in all ecosystems, as decomposers, symbionts of animals and plants and as parasites" (Dix & Webster, 1995, preface). By analogy, the monument to lost data will feed on the decaying matter of our information age, transforming it into something that might be more beneficial to the society. "Mycelium" is also attractive for its homophonic suggestiveness, namely its resemblance to the words "mausoleum" and "museum," which are also repositories for valuable things. The mycelium works well as the poetic structure for the monument to lost data because mycelia permit novel hybrids (see Figure 7). In the monument I envision, there will be novel combinations of data loss entries that will lead to new insights about the costs of our collective values and behaviors.

I envision the various data loss entries in the monument being linked using the principles of "hyphal fusion," a principle of linkage that fungi employ to share resources. Hyphal fusion permits diverse species of fungi to bond their mycelia together and share resources, and it also leads to the development of new species of fungi (Dix & Webster, 1995, pp. 15–16). By analogy, I use the principle of hyphal fusion as a poetics for structuring the virtual part of the monument to lost data; diverse data loss entries will be joined and thus will produce, I hope, new insights and possibly new memes for understanding our personal and

Figure 6: Proof of the concept — a monument to lost data, in the form of mushrooms, at the NARA in Washington, DC. As the NARA loses more data, more mushrooms will grow inside the building to represent the data loss (NARA image with mushrooms added by the author).

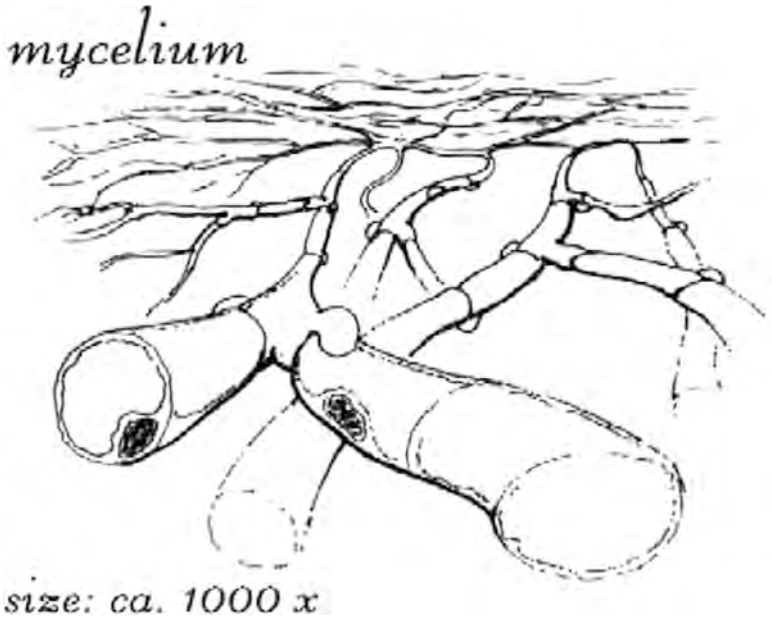

Figure 7: Branching network structure of the mycelium (http://home.wanadoo.nl/abiemans/pict/mycelium.jpg).

collective relationships to data. Such linkages do not impede users from exploring the monument's database by using a search engine.

By placing data loss entries into new relationships with other data loss entries, we draw each one into a new context and transform personal losses into sacrifices that can be recognized as parts of collective losses. We make a similar cognitive move when we see each soldier's death not merely as a personal loss, but as part of a larger context — a collective loss. If we were to juxtapose soldiers' deaths to other kinds of losses, such as the growing loss of the earth's ozone layer, we could search for meaning in the linkage. We might conclude, for instance, that the loss of soldiers in Iraq and the loss of the ozone layer are both connected to our consumption of fossil fuels. If that is our conclusion, then we might reconsider whether our consumption of fossil fuels is worth the costs.

The principle of hyphal fusion in fungi can be carried over to the design principles of the monument to lost data. The rhizome model put forth by post-structuralist theorists such as Deleuze and Guatari explains how the principle of linking diverse texts may serve as not only an archival structure in electronic databases but also as a means of conducting further research. Ulmer puts these post-structuralist insights into practice in his development of the puncept.

Ulmer imagines teaching a "new mimesis ... based on homophonic resemblance" in which the "puncept" would coincide with the "concept." Such a pedagogy, derived from "the fully developed homonymic program at work in Derrida's style" and amply evident in the present critical writing, would seize on the puns or homophones as precisely the device which, at the level of language, is "capable of relating elements with the least motivation, hence with the greatest economy of speed" and "generating the greatest 'information' (i.e., negentropy)." Such synthetic terms may offer an instance of the "higher-order bootstraps" students will need to think a massive cultural inventory soon accessible by the terabyte, in hypertext: as the neurologist Gerald Edelman observes, "[t]hinking occurs in terms of synthesized patterns, not logic, and for this reason, it may always exceed in its reach syntactical, or mechanical relationships" (Hilton, 1995).

The puncept will be the textual version of "hyphal fusion," or linkage, in the virtual part of the monument to lost data. It permits patterns to form based on shared terms and can be used in addition to the hierarchical groupings, like the Dewey Decimal System, found in most data archives. The monument to lost data will take terms "out of context" the way a Google search does, without regard for the topics of the texts being searched; thus the word "trace" may appear both in scientific works and in art criticism, and can be used to link texts from both discourses. The monument will treat terms that cross discourse boundaries as points of linkage among diverse entries and will present juxtapositions of such entries in order to help people produce new inferences about their data losses.

The first step to employing the puncept for hyphal fusion is to find an aesthetic pattern (the sharing of a common term or attribute) then move to find other patterns (of correlation, cause/effect, etc.). For example: years ago I assigned a monument project to my class at the University of Florida. One student produced a monument to losses caused by sickle-cell anemia. I made a monument about prisons. The connection was that discourses around both topics employed the word "cell," although the term meant something different in each discourse: blood cell in one and prison cell in another. Beyond that signifier linkage, we discovered that African–Americans found themselves in prison cells and with sickle cells

disproportionately more than whites. Another result of the "probing" that we did of the sickle-cell anemia/prison cell link was our discovery that social expenditures on these issues were disproportionate. In the case of prisons, disproportionately more money was spent sending black men to prison as was spent sending them to college. In the case of sickle-cell anemia, a lesser portion was spent on researching this disease compared to what had been spent on other diseases that affected whites in greater numbers. Such inferences gave us impetus to conduct additional research into race as a vital component linking both problems.

The idea of using the mycelium as a metaphor for the crossing of discourse boundaries is not new. In fact, Freud used this same metaphor over a hundred years ago to explain the way that discourses in dreams formed networks around a discursive "navel." As Weber pointed out, Freud's use of the mycelium metaphor, rather than the more common "branching roots" metaphor, created interesting problems for Freud precisely because, unlike the branching roots structure, the mycelium is a decentered structure.

> The "navel of the dream" is read through "dem Unerkannten aufsitzt" by Weber as "an untenable alternative" which Freud "straddles" (81): what I call, following Derrida, undecidability. The reference to botany leads Weber to focus on the last line of the passage — "Out of one of the denser places in this meshwork, the dream-wish rises [erhebt sich] like a mushroom out of its mycelium" — and his next piece of evidence that Freud ... is interested in the dream as dislocation, "ent-stellt" (81). Referring ... to the O.E.D., Weber finds the following for "mycelium": "Mycelium. (f. Gr. mykes mushroom, after epithelium) Bot. The vegetative part of the thallus of fungi, consisting of white filamentous tubes (hyphae); the spawn of mushrooms" (qtd. in Weber, 81). And for "thallus": "(Gr. thallos, green shoot, f. thállein to bloom) Bot. A vegetable structure without vascular tissue, in which there is no differentiation into stem and leaves, and from which true roots are absent" (ibid.) ...
>
> Interpretation's straddling of the dream-navel, and the thallus (non-roots) at the root or origin of the dream (the dream-wish), both suggest the undecidability of an unknowable rather than an unknown. This movement and undecidability, for Weber, suggest the unknowable, a différance and not a specific absence of meaning: a non-original mobile textuality rather than an original immobile text. (Anders, 2000)

The monument to lost data, like the meshwork of the dream, ought to be decentered, supporting an organic process of cultural invention. The individual participants — the mourners who add their traces and phantoms to the monument — will see their contributions linked to others. As they visit the monument and see the links to and from their entries, they may come to understand the collective dimension of their personal data loss. Thus the loss of a parent's living memory to Alzheimer's might be understood in relation to the loss of a hard drive. Similarly, the collective dimensions of data loss might be re-imagined as part of a global shift in data-storage practices. The loss of native languages might be understood in relation to the loss of Kosovars' identities, following the link between the

Kosovars' stolen passports and the minimal amount of porting (a computer networking term) of native languages onto the Internet.

Could the design I propose here lead to relationships that trivialize great losses (as might be the case if there were to be a link between the loss of a language and the loss of an iPod collection of pop songs)? To some extent, there is a risk of trivializing great losses by linking them to lesser ones. But there are also advantages to making all kinds of juxtapositions and relationships without first considering relative degrees of loss. To some extent, we can distinguish between the loss of memes and personal losses, but that line may blur; what if thousands of people were to lose their iPod collections? Would those losses be trivial? A hyphal node juxtaposing the relationship between the loss of a language and the loss of an iPod collection might prompt the iPod owner to reconsider the value of her loss. The monument to lost data itself may appear to trivialize other monuments, such as war monuments. Such appearances might give us cause for reflection, but should not deter us; the potential benefits of the monument to lost data are too great.

The monument to lost data should not be a fixed entity; its components should be linked in a process that continually renews the monument. In the execution of the monument to lost data, numerous problems will undoubtedly arise. For instance, who will make decisions about filtering content in the monument? What criteria will be used to make such decisions? Which entries in the monument will receive prominent placement, or will all placement be equal? These questions have been raised in relation to a number of public and semi-public projects, including the Names Project and Wikipedia. Large-scale public and semi-public databases are vulnerable to submissions that might be perceived as offensive, controversial, or trivial. I would argue, however, that the monument to lost data should take as broad and egalitarian a view as possible toward possible submissions. Consider a hypothetical problem that could arise; one group memorializes the loss of Indian sacred lands while another group memorializes the disappearance of statues of Columbus. Without judging the relative merits of these submissions, the monument's managers — a group that may be entrusted to guide participants' behaviors the way that web moderators do — might provide a setting and guidelines for the participants to engage in dialogue with each other. The best practices for such moderation will emerge from experience and good judgment as the project comes to fruition.