

Digital media, and the coming gap in the historical record

During the early days of the internet, web proponents were quick to make expansive claims about the benefits that the on-line age would usher in. The web would revolutionize the work environment, they said. It would engender a new era of political activism. It would preserve all human knowledge on-line, forever.

Some of these claims have turned out to be quite prescient. Others are still open to debate. And some linger in the murky twilight of being true, but only partly so. The promise of the web as the ultimate archival medium falls into this last category.

While the internet contains an incredible amount of information, the life-cycle of web-based material is anything but certain. The web preservation specialists at the Internet Archive estimate that the average life-span of a web page is a mere 45-75 days, with the deleted data quickly escaping into the nether-regions of cyberspace. Some of this information is lost through periodic site updates, and some through the inevitable attrition of websites themselves. Particularly in the highly fluid world of personal blogs, once an owner loses interest and neglects to pay the annual hosting fee, the site eventually disappears, to be replaced by the omni-present grave marker of the internet age – the “404 error.” Even the most diligent website owner is hampered by the ultimate impediment of their own mortality. Beyond that horizon, no site can survive without outside support.

A challenge for history

The fragile nature of the web should concern us all, because of the internet’s increasingly central role in disseminating knowledge, and its parallel role as the host for a vast amount of original, internet-only information. The web’s centrality, coupled with its impermanence, presents a unique and pressing dilemma – the specter of valuable cultural information rapidly disappearing down the memory hole.

Because of its vibrant on-line culture, Minnesota is particularly susceptible to this problem. In recent years, our state has produced a raft of prolific on-line authors of all ideological stripes and personal backgrounds. This has also been the case with web-based news reporters and public affairs

commentators. Minnesota has been an innovator in the creation of all-digital news outlets, and our state's plethora of sites frequently cover stories that are not duplicated by other mediums.

A pressing question that is raised by the proliferation of these on-line venues is one of archival integrity. As more and more of Minnesota's public conversation goes digital, what steps should be taken to ensure that history will have access to it?

Problem not unique to the web

While the web poses distinct archival dilemmas, the issue of digital impermanence is not unique to the internet alone. Those working in television and radio have long been aware of the fragile nature of the mediums they toil in, and recent technological updates have raised new difficulties for them, similar to those that face the internet. It's worth examining these problems briefly for the sake of comparison.

In television news, for instance, the switch to digital media has brought many short-term benefits, but it has also complicated archival preservation. While lowering the cost of production and increasing the ease of dissemination, digital video is highly transitory, and subject to disappear without careful planning. In contrast, motion picture film - the original medium used for location news gathering - was quite expensive, but it had a projected shelf-life of well over a century.

More recently, magnetic tapes replaced film-based capture in most news departments, bringing an increase in ease, but a decrease in longevity. These tape formats changed rapidly, and some of the earliest recordings have already suffered from serious deterioration. This is likewise true of magnetic audio recordings.

The digital video revolution of the 1990s sped up the proliferation of videotape formats, making retrieval ever more dependent upon highly specific equipment. Now, with the advent of tapeless digital recording, video material relies on continual hard drive back-ups - and frequent software updates - in order to keep it accessible over time. Digital audio recordings face similar issues.

Television and radio material has always held problems for archivists, due to its fragile and highly technical nature. However, past preservation

efforts were always able to rely on tangible artifacts. Not so today. As video and audio programs go entirely digital, these mediums are starting to parallel the internet, in that they are reaching a point where there is no underlying artifact to preserve.

Archival disparities

For large institutions, the task of preserving substantial volumes of digital information is daunting, but not impossible. Minnesota Public Radio, for example, is currently in the process of creating a digital archive of all of its magnetic tape recordings, ushering in a future of massive, ongoing data management. For individuals and small organizations that do not have the resources to invest in such long-term solutions, the digital age may portend an archival black hole.

For instance, 8mm home movies can sit in your uncle's attic and still be recovered in 80 years, but can the same be said of the 720p HD Quicktime files stored on your daughter's laptop? Indeed, will her laptop even function in a decade? Posed this way, the preservation predicaments of the digital age quickly become apparent.

Archiving the internet

On the web front, the problems are related, but slightly different. As a society, we face an uncertain and changeable digital future – one in which some other delivery system will likely supplant the internet as we now know it. At the very least, some other standard is bound to eclipse the current HTML website design language, potentially making older web material unreadable at some future point. Both of these scenarios pose real dilemmas for today's trove of original, web-only information.

Because of these problems, efforts are now underway to preserve web-based information for future reference. For example, the San Francisco-based Internet Archive has partnered with major institutions like the Smithsonian to help save the internet's digital artifacts. The Internet Archive has been at work for several years now, "crawling" the web with software designed to take snapshots of public web sites. The Archive then makes these snapshots available in an on-line portal called the "Wayback Machine."

By sifting through the Wayback Machine's search engine, one can find remnants of dead sites, as well as early versions of current sites. Many of these sites are missing critical bits of information, however. For better

preservation results, the Internet Archive suggests that website owners register their sites with them directly, in order to allow the Archive to compile site changes more frequently and comprehensively.

Small institutions and individuals should be aware of these efforts, and should make sure to submit their websites for inclusion in this on-line preservation enterprise. However, a more traditional option is also available to those concerned with archival longevity.

A paper future?

In preservation, redundancy matters. While there has been a trend by museums to place digital copies of their collections on-line, these institutions are still retaining and preserving original, physical artifacts. This is particularly true of paper. After all, high-quality paper, properly preserved, is an incredibly stable medium. We need look no further than the Dead Sea scrolls or the Declaration of Independence for proof of this.

Even in the midst of the digital explosion, one should not doubt the central role of traditional archival institutions (and their hard-copy collections) in the formulation of the historical record. The printed news story is still the single most important record for writers and researchers looking to re-assemble the bits of history's ragged narrative. The personal, written correspondence of soldiers, politicians, and citizens also provides valuable context for historians seeking to understand the reasons and passions that drove historical events.

Today, blogs serve a similar function to the personal letters of years past. They chronicle individual responses to social events, and give us a more comprehensive picture of our era's psychology. Today's proliferation of on-line journalists and media producers also offer original perspectives not available elsewhere, and their materials hold unique historical value. This is why it is important for those working without large-scale institutional support to preserve their work in the most future-proof format possible. In short, that format is paper.

Independent journalists and commentators: make sure to transcribe your work and commit it to the printed page. On-line writers and bloggers: print the contents of your sites comprehensively. Save a copy in your attic, send one to the copyright office, and most importantly, donate a copy to your local or state historical society. History will be glad that you did.

Originally published at Minnpost.com | April 3, 2009