All things must pass – except maybe your tweets

Immortality and the Internet

"Nothing posted ever really disappears" has already become a nugget of proverbial lore, quoted by wise elders to the young to warn them against sharing potentially embarrassing information online. But like most chunks of received wisdom, it's not strictly true.

As we all learn on the playground, unpleasant facts about us seem to last forever and can pursue us for far too long. If such things now spread virally, they can inflict damage for a seeming eternity, but is it really possible for data to last *forever* on the Internet?

How long do webpages, and even memes, endure? Can data be preserved indefinitely? And what happens when the people who generate information pass away? Can there be life after death on the Net?

This issue will examine the question of life expectancy of information on the Internet. Now that the Web is over a quarter of a century old, the issue of **data preservation** is becoming a serious concern not just for old folks online, but for historians, archivists, statisticians, and record keepers of all kinds. And it has significant implications for the future too.

Digital afterlife

Some people dream of a final state where their personality will be transferred into more durable media than frail flesh. Such **mind uploading**, even if possible, would at best be a computer-generated model that mimics the deceased. However, Google has already patented a system whereby artificial digital personalities, maybe including a simulation of your dear Aunt Sally, could be **downloaded into robots**.

But, like those gone before, we'll likely have to settle for memorials, and as other aspects of our lives have moved online, they will be built in cyberspace. It's generally been up to friends and family to pay for setting up tributes or preserving a deceased person's website. However, **Facebook**, as the town square of the digital age, has since 2009 allowed **digital memorials** which would preserve a few pages and allow friends to share their memories and grief. Even private **memorial groups** can also be set up.

Sites like **Eternime** promise people the chance to be "virtually immortal" by offering users a way to set up their own digital memorial. Unfortunately, they offer few details, which is a concern as a similar start-up, Intellitar, got 10,000 people to sign up to create online avatars (for $25/month) – and then **folded.**

Another start-up, **DeadSocial**, offers an app allowing users to post news of their own passing to social media. However, all these schemes depend on the sites being there in the future. But there's no guarantee that even Facebook and Google will last forever.

Building on quicksand

The Internet is a work constantly in progress. It was invented as a means of transferring data, not preserving it. Yet with so much of modern life dependent upon its workings, the Net has been forced into assuming the role of the ultimate **data repository**, an oracle with access to everything ever posted.

But in actuality, the Web and all its contents are constantly changing – and becoming ever more transient and complex. The **lifespan** of a static webpage was estimated in 1997 to be just *44* days; a study in 2001 indicated a life of *75* days, and one in 2003 gave *100* days. But there's no assurance whatsover that the information on that page hasn't changed radically any number of times even in that brief period.

Hyperlinks, the essential tags that bind all the information on the Web together, are even more fleeting.

Link rot is universal, affecting everything from legal and scholarly citations and journalism to school papers. Though resources like Wikipedia **suggest** means of limiting link rot, many references on their own pages become unusable in just a few months.

The reasons are varied: websites can be restructured, renamed, or die; access policies may change (as often happens with news sites); or the data itself might be very impermanent in nature. This is particularly true of dynamic pages, like those in social media, where constantly developing practices may change what can be seen by whom or which foil searches.

Along with link rot comes **software rot**. Programs evolve, changing formats, and old software dies. Software, like **Adobe Flash,** that becomes a steady staple in one era may be widely abandoned later on.

The biggest danger seems to be **irrelevance**. Data is saved as long as it's felt to be important. Lost silent **films** and early episodes of *Doctor Who* vanished simply because people thought shelf-space or video tapes were more valuable than the shows on them.

The usual response is to copy and stash material. When Yahoo shut down its once-popular **Geocities websites** in 2009, fans rallied to preserve as much as *900 GB* of the early, and often crude and ugly, pages.

They were aided in this by the **Internet Archive**. Since 1996, under the slogan "universal access to all knowledge", the **Archive** has been busily collecting billions of digitized materials. They range from public domain library book collections to movies, cartoons, videos, music, games, and websites. Access is free, though they have a subscription-based service, **Archive-it,** to help libraries and museums build their own. They also run the **Wayback Machine**, using web crawlers to capture websites, visiting every few weeks or months, now with *439 billion* pages saved.

This is still far less than the **nearly *one billion* web-sites** out there, and even the founders doubt that *everything* can be captured, due to broken links, rapidly-changing dynamic pages, and missing data – not to mention legal restrictions and an increasingly fractured Web. Despite their best efforts, they **admit** that the original Web of the 1990s simply "does not exist".

The coming digital Dark Age

The task would be much easier if we already knew what the future would consider important, of course. Failing that, everything must be considered fair game – even tweets. The **Library of Congress** is building an enormous Twitter archive, while **Google**, **Bing** and others are allowing searches for even the most inane comments, though you can still **delete** yours.

Data stored in the cloud has a better chance of enduring than in physical media, because it can last as long as it is transferred from one server to another in time. For along with evolutionary changes in data formats and the **planned obsolescence** of hardware that threaten the usefulness of storage media, such backups themselves are all too fallible.

In a **previous issue**, we looked at the short, uneven lifespans of memory media. Hard drives *might* possibly last up to *34* years, but flash drives and CDs, just *3-10* years. Though developments of **crystal memory** technologies that last almost forever and hold astronomical amounts of data have been promised, they have not made it out of the lab yet.

Even initiatives like the Library of Congress' seeking **new ways** to preserve news for future researchers might not be entirely successful. Some scholars even speak grimly of a "**digital Dark Age**" where knowledge of this vital period of transition will not exist.

If we're lucky, humanity is passing, kicking and screaming, from the Industrial Age of warring nation states to an Information Age and a more peaceful, united planet. Unless, of course, disasters likely of our own devising plunge us headfirst into the brutal darkness again. Whatever happens, our heirs will want and need to know how it came about.

But most knowledge will be in digital format, and the great problem with electronic information is that when it's gone, it's *really* gone. For by its very nature, such data is as fleeting as a spark. Even if it can be preserved in memory banks, what happens *after* the power goes out? Today even **charred texts** can be read with the right tools but no similar means yet exist for long-dead electronic data recovery.

The loss to future historians of our online trivia might not be that bad, but to lose *everything* would be an unthinkable catastrophe. Scholars estimate that the epic tales of the Trojan War once filled as many books as the Harry Potter series. All we have left of them are most of the first and last, with no more than a few scraps in between. The rest went missing in the last Dark Age. We owe it to our grandchildren's grandchildren to prevent such tragic losses from repeating.