Personal Archiving and Preservation
through Social Media

Lisa Peck
INFO 601-05
12/10/19

The personal archive is as old as time itself—a collection of personal effects that allows us to research a person, and learn about their lives from the way they interacted with and saw the world. Personal archiving has evolved with time and technological advances, from letters and scrapbooks, to photographs to audio and video recordings. Many personal archives are made of matter intentionally saved and curated by their subject, whether this be research papers or something more personal, such as scrapbooks.[1]

For many researchers, these personal archives are invaluable. However, as we live in a continuously mediated digital world, electronic forms of communication have all but replaced handwritten correspondence. Digital photos and videos rule supreme. Traditional ephemera has been eclipsed by its digital counterpart. How are we to do personal research in the future if these ephemeral documents are no longer as abundant?

We must rethink our current definition of the personal archive, and look to the ways that people have begun to document their lives digitally. For many people, this means utilizing the Internet. Social media is one of the largest forms of social interaction; billions of people worldwide use some form of social media, documenting their lives online daily.

In this paper, I will be discussing Internet usage by the average user of social media rather than online personalities. "Influencers" interact with the online ecosystem in a way that is not representative of the majority of those using social media. People that have built a career online use the documentation of their lives on social media as a tool to further their business. I will reflect on influencers when appropriate, but strive to focus on the way that an average person uses social media to represent themselves. Additionally, when referencing "user data" in

---

[1] Katie D. Good, "From scrapbook to Facebook: A history of personal media assemblage and archives," *New Media & Society* 15, no. 4 (2012), doi:10.1177/1461444812458432.

this paper, I am referring to the content created, posted, and shared by people, not usage data collected by various platforms.

Personal Archiving…. Online?

When one considers personal archiving in digital form, the first thing that comes to mind may be a personal journal or blog. Blogging is perhaps one of the oldest forms of communication on the web. Early adopters kept personal homepages, detailing their lives in long form. These homepages gave way to blogging platforms for those with less technical knowledge. Specialized sites like Wordpress, LiveJournal and Tumblr allow their users to create blogs in which they chronicle their lives, supplemented with photos, videos, and whatever else the writer chooses. Sites encourage the tagging and sharing of posts, allowing one's journal to be publicly searchable. This flexibility not allowed by the print medium has created an explosion of content. Blogging is one of the most prevalent mediums of communication on the web, as writers may easily and freely self publish on any topic they choose.

One such popular topic is family and lifestyle blogging. Users write about the minutiae of their daily lives, and the strategies (and often sponsored products) that have helped them through these experiences. Family bloggers (nicknamed mommy bloggers) are known to document in detail their child's lives, sometimes beginning in utero. This phenomenon is akin to keeping a digital baby book - a highly detailed record of baby's first smile, haircut, milestones. These blogs exist perhaps for the benefit of distant family and followers, but also as a personal record for the writer.

Most looking to connect with family use what is perhaps the most ubiquitous and well known social networking site, Facebook, which boasts an estimated 2.14 billion active users

monthly.[2] Facebook usage begins with the creation of a profile using a first and last name, asking people to use their real identity online. Anonymity is discouraged on Facebook, as the site revolves around genuine connection with family and friends. The site asks individuals to add personal information about themselves, such as birthday, hometown, relationship status, and employment history. Most willingly add this information, updating their profiles with every major life change.

Facebook prompts, "What's on your mind?", asking you to share your thoughts, photos, videos–the options are seemingly endless. What is the motivation to post about your life online? For many, it is the audience. Often, Facebook friends span years; the people you are connected with on Facebook are not necessarily those you are currently closest with. The desire to share is tied to the very human desire to put one's best face forward, and the website's sharing structure encourages this.

Early Facebook profiles featured every action taken by the account mixed in with personal content such as photos, status updates, and wall posts. In 2011, Facebook announced a new layout for personal profiles known as the 'timeline', writing: "Now, you and your friends will finally be able to tell all the different parts of your story – from the small things you do each day to your biggest moments."[3] Rather than an unstructured list of posts, the new timeline layout highlights important content, and enables one to select a specific year, or even month, to look back at and reminisce. Though the average person may not have seen posting status updates as a way of archiving their life, it is clear that with the creation of the timeline layout, Facebook

---

2 "Facebook Users Worldwide 2019," Statista, accessed November 16, 2019, https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/.

3 Sam Lessin, "Tell Your Story with Timeline," Facebook, accessed November 16, 2019, https://www.facebook.com/note.php?note_id=10150289612087131.

recognized this possibility, encouraging their users to "collect all [their] best moments in a single place" utilizing life events.[4]

Life events allow people to commemorate occasions, ranging from the more mundane (Braces removed!) to the life-changing (Expecting a baby!). Facebook automatically adds certain life events, such as moves and job changes, based on edits you make to your profile's basic information. However, additional events, such as "Became an organ donor," have to be sought out. Some even choose to back-date these events, adding significant moments that occurred in their life before Facebook was even invented.

The metadata one can input about their own life is ever-expanding. In 2013, Facebook gave their users the option to input "Professional Skills," moving beyond life events and personal history into a more fully-rounded professional and objective realm.[5] Traditionally, Facebook has focused on documenting one's personal life. The move into a more professional setting–Facebook has also since integrated job listings and applications onto its site–suggests that Facebook is truly interested in collecting and archiving a complete snapshot of each person–from their hobbies and interests to their proficiency level in Excel.

It is worth noting that while the average person may not go into extreme depth about their life on Facebook or other social media platforms, many choose to upload at least some information about their lives. Even the most minimal data creates an unintentional archive, much like a historical census record.

---

[4] ibid.

[5] Adario Strange, "Facebook Adds Professional Skills Category for Job Seekers," Mashable, last modified September 8, 2013, https://mashable.com/2013/09/08/facebook-jobs-skills/.

YouTube is the largest video host online–the site revealed in May 2019 that over 500 hours of new footage is uploaded every minute.[6] While a large amount of this traffic comes from channels run by content creators and corporate entities for whom YouTube is a job, a large portion of content is generated by average people. Free and reliable, YouTube is the universal go-to site for sharing footage with others.

Individuals may upload acting reels, sports highlights, school performances, student films, or videos of their families and friends to share. These videos, particularly those of performances, are often linked to the subject's real name, allowing for ease of findability. As people add more personal information and metadata to social media to flesh out their profiles, their content becomes more and more discoverable.

There are countless websites and apps available for people to archive their lives in any way they choose, whether they are doing so knowingly and intentionally or not. On many social media sites, keywords and hashtags aid in the discoverability of specific posts, and can be used to find other similar posts, furthering discussion. However, one can also use tags in the opposite way; using hyper-specific hashtags, creating a virtual timeline of your own posts.

As a personal example, one of my cousins has two children. With each photo she posts of them online, she adds a hashtag comprised of their full name, affixed with their current age. A public search of these hyper-specific hashtags reveals the photos of each child at a certain age in the order of their posting, allowing for (I can only assume) her own reference and reminiscing.

Another example of this phenomenon is wedding hashtags–a trend in which a couple chooses a hyper-specific hashtag for the event (often a pun based on their combined last names)

6 "YouTube: Hours of Video Uploaded Every Minute 2019," Statista, accessed November 18, 2019, https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/.

and encourages guests to use the hashtag when posting photos from the wedding. This allows the couple to search for their personal hashtag, revealing all photos posted that they may have missed. This practice allows them to easily look back on their wedding from the guest's perspective. However, this practice also assumes that the chosen hashtag is in fact, unique. Sometimes wedding hashtags can overlap, painting a murky picture of the events. Hyper-specific hashtags are imprecise, and only useful on one platform at a time.

Preserving Social Media Archives

While we create and curate these online archives, we seldom think about their digital lifespan. Growing up, my peers and I were so often told that "everything on the Internet lasts forever" that I find we are now lackadaisical about making a concerted effort to save our own content. If we share freely about ourselves and our lives online, what happens to this information after our death? Traditional paper based personal archives are often formed after one's death, from materials left behind, or specifically saved by the deceased. Are our digital archives not the same?

Though most only offer deletion, some prominent social media sites offer specific services upon a person's death. For example: Instagram offers a nearly invisible "memorialization" option, allowing a profile and content within to remain static as-is after death, and removing the deceased from the "Explore" tab.[7] Google allows a trusted individual to download another user's data after a "period of inactivity," meaning death or another

---

[7] "What Happens when a Deceased Person's Account is Memorialized?" Instagram Help Center, accessed December 10, 2019, https://help.instagram.com/231764660354188.

incapacitating event.[8] Perhaps unsurprisingly, the most comprehensive of these memorial options is offered by Facebook.

If a person chooses to archive their account after their death, Facebook will close the account, leaving posts and profile photos to remain online in the state the user left it indefinitely. The title "Remembering" is added to the profile, and friends are able to continue to interact with the deceased's page and the content as it was at the time of the page's memorialization. Furthermore, the deceased will not appear as a suggested attendee for events and other interactions, limiting the possibility of painful reminders of their passing.[9] When this feature was announced, many found it uncomfortable to think about their digital lives after death. Personally, I believe that Facebook is on the forefront of a necessary discussion.

Notably, Facebook's policy, was well as that of many other social media sites, requires proof of death from a surviving family member or friend. Facebook encourages its users to select a "legacy contact" - one who will decide on behalf of their friend whether to archive or delete their Facebook profile upon their death. However, without significant and appropriate forethought, the agency to make their own choices about their digital legacy is lost.

Trusting these platforms to maintain and preserve our personal data for posterity is a tempting choice. After all, Facebook is often used as a forum for grieving individuals to find solace and community with other mourners.[10] However, what happens to the data saved on memorialized profiles when the website hosting it shuts down?

[8] "About Inactive Account Manager," Google Help, accessed December 10, 2019, https://support.google.com/accounts/answer/3036546.

[9] "About Memorialized Accounts," Facebook Help Center, accessed December 10, 2019, https://www.facebook.com/help/1017717331640041/?helpref=hc_fnav.

[10] Anna Haverinen, "Facebook, Ritual and Community – Memorialising in Social Media," *Ethnologia Fennica* 42 (2016), doi:10.23991/ef.v42i0.59284.

A veritable giant in the tech world, it is a safe bet that Facebook and its subsidiaries will exist well into the future. Historically, reliance on third party platforms to maintain and archive user data indefinitely is notoriously dangerous; MySpace recently accidentally lost decades of data during a server move.[11] Sites like GeoCities, a popular host for early web homepages, have been lost to the world by decision of its parent company Yahoo!, and contemporaries, like popular file storage site Dropbox, routinely undergo intentional systemic purging of inactive accounts and their data.[12]

Self published sites, while offering complete control over published content, are not immune from virtual decomposition; these sites are potentially even more vulnerable as they must be maintained by their owners or a third party. Even still, outdated technology used to create these pages and broken links lead to loss of content.[13]

Virtual activist groups are striving to save all site and user data before it's too late. One such activist group, Archive Team, maintains a watch list of significant sites in decline, and important sites that serve as hubs of information that must be carefully monitored for signs of shutdown.[14] These tell-tale signs include recent changes in site management, such as sudden lay-offs and acquisitions. As smaller sites may be more vulnerable to data loss, they are more closely monitored. The larger and more active the user base, the more unlikely it is that the site will be

---

[11] Niraj Chokshi, "Myspace, Once the King of Social Networks, Lost Years of Data From Its Heyday," *The New York Times*, last modified March 19, 2019, https://www.nytimes.com/2019/03/19/business/myspace-user-data.html.

[12] Jordan Novet, "Dropbox Shows How It Manages Costs by Deleting Inactive Accounts," CNBC, last modified February 23, 2018, https://www.cnbc.com/2018/02/23/dropbox-shows-how-it-manages-costs-by-deleting-inactive-accounts.html.

[13] Google Chrome will not support Flash after December 31, 2020, making countless sites from the early web inaccessible to those using the browser. At my current place of employment, we host dozens of educational resource sites created by our office in the early 2000's that are Flash based and too time consuming to upgrade.

[14] "Archiveteam," Archiveteam, accessed November 19, 2019, https://www.archiveteam.org/index.php?title=Main_Page.

shut down. However, the site's size alone does not mean that user data is automatically safe from systemic deletion.

Much of the web's data is interwoven, dependent on each other to survive. Many social media users rely on content others have posted (ie. tagged photos) to make up their profile data. This begs the question: what happens to our media when access disappears?

In 2016, in a stand against corporate interference with open data, a programmer removed all of his open-source code packages from npm, a site facilitating open JavaScript software exchange. One of these packages, *left-pad.js*, was referenced in numerous other packages, and this sudden deletion had created dependency errors in thousands of sites.[15] Many affected developers were completely unaware that plugins included on their sites depended on *left-pad.js*, leaving the JavaScript community floundering for answers.

On a smaller, more personal scale, this type of unexpected loss is a model for what we risk by leaving our content solely in the hands of platforms, or by relying only on content posted by others for our own archive. We must make the effort not only to save our own data, but to save the contextual information surrounding it whenever possible. As so much of the open Internet is interconnected by hyperlinks and references that even if textual content is saved, it is for naught if the content linked within has disappeared.

In 2003, Roy Rosenzweig wrote about the urgent need for digital preservation that has existed since the early 1980's. Rozenweig writes, "preserving the born-digital materials for future historians seems like a theoretical and technical issue, tomorrow's problem," illuminating

---

[15] Keith Collins, "How One Programmer Broke the Internet by Deleting a Tiny Piece of Code," Quartz, last modified March 27, 2016, https://qz.com/646467/how-one-programmer-broke-the-Internet-by-deleting-a-tiny-piece-of-code/.

a mindset that has held back digital preservation strategies.[16] Given how rapidly technology has advanced and Internet users have expanded in the last 16 years, the need for an overarching comprehensive digital preservation strategy has only increased in urgency, yet there is none to be found.

Independent Internet archival groups are overwhelmed by data; Archive Team selectively downloads user data of sites at risk, whereas the Internet Archive strives to collect as much of the open web as possible through random site selection.[17] The Library of Congress once began an effort to preserve every single tweet posted on Twitter from its inception in 2006. As Twitter grew, this collection strategy became untenable, and LoC instead chose to archive only selected tweets after 2017.[18] Similar efforts to archive an entire site's worth of user data have not been undertaken by US governmental agencies since.

Library of Congress has stated that it will continue to archive selected tweets, one such Twitter account will more than likely be that of the President of the United states, @POTUS. In 2014, the Presidential Records Act, which dictates presidential record preservation, was amended to "include electronic content, … interpreted as an umbrella term encompassing text messages, emails, social media and the like." [19] After his term in office, President Obama's tweets made

[16] Roy Rosenzweig, "Scarcity or Abundance? Preserving the Past in a Digital Era," *The American Historical Review* 108, no. 3 (2003), doi:10.1086/529596.

[17] Internet Archive: Digital Library of Free & Borrowable Books, Movies, Music & Wayback Machine, accessed November 19, 2019, https://archive.org/.

[18] Elisabeth Fondren and Meghan Menard McCune, "Archiving and Preserving Social Media at the Library of Congress: Institutional and Cultural Challenges to Build a Twitter Archive," *Preservation, Digital Technology & Culture* 47, no. 2 (2018), doi:10.1515/pdtc-2018-0011.

[19] Rachel Treisman, "As President Trump Tweets And Deletes, The Historical Record Takes Shape," NPR, last modified October 25, 2019, https://www.npr.org/2019/10/25/772325133/as-president-trump-tweets-and-deletes-the-historical-record-takes-shape.

while in office were archived and are preserved by the National Archives and Records Administration.

It can be argued that the content posted by the average person is less important than that of significant thought leaders, politicians, and the like. However, it is also arguably more important to preserve as much of the cultural fabric of the Internet as possible, especially that of the average person. Activist groups are prepared to help; in addition to archiving other's data, Archive Team offers information for people looking to save their own data. Many platforms, notably, Facebook, Twitter, Instagram, and Google, offer official tools to allow users to download their data. These files contain the content uploaded by the individual was well as data collected by each platform about the account.

Social media provides the ability to post about our daily lives, creating a digital representation of ourselves in the process. These sites act as a ubiquitous personal archive for many people, and must be preserved as such. It is not feasible or realistic to expect corporations or the government to preserve our social media data. It is imperative for people using social media sites to archive their own data, reducing the critical conservation burden of groups like Archive Team, and ensuring preservation of content for future generations. If we as users and as people are serious about harnessing the power of social media as an archive, we must in turn be serious about its preservation and seizing ownership over our own digital legacy.

Bibliography

"About Inactive Account Manager." Google Help. Accessed December 10, 2019. https://support.google.com/accounts/answer/3036546.

"About Memorialized Accounts." Facebook Help Center. Accessed December 10, 2019. https://www.facebook.com/help/1017717331640041/?helpref=hc_fnav.

"Archiveteam." Archiveteam. Accessed November 19, 2019. https://www.archiveteam.org/index.php?title=Main_Page.

Chokshi, Niraj. "Myspace, Once the King of Social Networks, Lost Years of Data From Its Heyday." *The New York Times.* Last modified March 19, 2019. https://www.nytimes.com/2019/03/19/business/myspace-user-data.html.

Collins, Keith. "How One Programmer Broke the Internet by Deleting a Tiny Piece of Code." Quartz. Last modified March 27, 2016. https://qz.com/646467/how-one-programmer-broke-the-internet-by-deleting-a-tiny-piece-of-code/.

"Facebook Users Worldwide 2019." Statista. Accessed November 16, 2019. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/.

Fondren, Elisabeth, and Meghan Menard McCune. "Archiving and Preserving Social Media at the Library of Congress: Institutional and Cultural Challenges to Build a Twitter Archive." *Preservation, Digital Technology & Culture* 47, no. 2 (2018), 33-44. doi:10.1515/pdtc-2018-0011.

Gehl, Robert. "YouTube as archive." *International Journal of Cultural Studies* 12, no. 1 (2009), 43-60. doi:10.1177/1367877908098854.

Good, Katie D. "From scrapbook to Facebook: A history of personal media assemblage and archives." *New Media & Society* 15, no. 4 (2012), 557-573. doi:10.1177/1461444812458432.

Haverinen, Anna. "Facebook, Ritual and Community – Memorialising in Social Media." *Ethnologia Fennica* 42 (2016). doi:10.23991/ef.v42i0.59284.

Internet Archive: Digital Library of Free & Borrowable Books, Movies, Music & Wayback Machine. Accessed November 19, 2019. https://archive.org/.

Lessin, Sam. "Tell Your Story with Timeline." Facebook. Accessed November 16, 2019. https://www.facebook.com/note.php?note_id=10150289612087131.

Marshall, Catherine C. "Biography, Ephemera, and the Future of Social Media Archiving." *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries - JCDL '18*, 2018. doi:10.1145/3197026.3197051.

Novet, Jordan. "Dropbox Shows How It Manages Costs by Deleting Inactive Accounts." CNBC. Last modified February 23, 2018. https://www.cnbc.com/2018/02/23/dropbox-shows-how-it-manages-costs-by-deleting-inactive-accounts.html.

O'Sullivan, Catherine. "Diaries, On-line Diaries, and the Future Loss to Archives; or, Blogs and the Blogging Bloggers Who Blog Them." *The American Archivist* 68, no. 1 (2005), 53-73. doi: 10.17723/aarc.68.1.7k7712167p6035vt.

Rosenzweig, Roy. "Scarcity or Abundance? Preserving the Past in a Digital Era." *The American Historical Review* 108, no. 3 (2003), 735-762. doi:10.1086/529596.

Sinn, Donghee, and Sue Y. Syn. "Personal documentation on a social network site: Facebook, a collection of moments from your life?" *Archival Science* 14, no. 2 (2013), 95-124. doi:10.1007/s10502-013-9208-7.

Strange, Adario. "Facebook Adds Professional Skills Category for Job Seekers." Mashable. Last modified September 8, 2013. https://mashable.com/2013/09/08/facebook-jobs-skills/.

Treisman, Rachel. "As President Trump Tweets And Deletes, The Historical Record Takes Shape." NPR. Last modified October 25, 2019. https://www.npr.org/2019/10/25/772325133/as-president-trump-tweets-and-deletes-the-historical-record-takes-shape.

"What Happens when a Deceased Person's Account is Memorialized?" Instagram Help Center. Accessed December 10, 2019. https://help.instagram.com/231764660354188.

"YouTube: Hours of Video Uploaded Every Minute 2019." Statista. Accessed November 18, 2019. https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/.