

The Lifetime Library Project: Facilitating a Lifetime of PIM

Daniel Beaver-Seitz

University of North Carolina – School of Information and Library Science

100 Manning Hall
Chapel Hill, NC 27599

dmbeaver@live.unc.edu

Terrell Russell

unc@terrellrussell.com

ABSTRACT

We describe the need for a tool for aggregation and curation of user-created digital artifacts, which would otherwise be scattered across multiple storage locations, and early work on such a tool.

Categories and Subject Descriptors

H.3.7 [Information Storage and Retrieval]: Digital Libraries – collections, systems issues, user issues.

General Terms

Design

Keywords

Personal digital libraries, personal archives, social media, reminiscing.

1. INTRODUCTION

The rise of user-generated content (UGC) has transformed both the quantity and quality of digital artifacts created by individuals. Individuals now create large corpora of digitized and born-digital works upon both general use platforms (e.g., Facebook) and more specific platforms, such as virtual learning environments (VLE), like Sakai or Blackboard. Unfortunately, these digital artifacts are often scattered across many platforms, creating a host of issues related to control and archiving.

We are developing a suite of open source tools with support for assembling unified collections otherwise normally stored at a variety of third-party Internet resources. By providing users with centralized storage space and automated harvesting tools, we aim to empower them to control their digital artifacts and create future value for those artifacts.

2. THE PROBLEM

The dream of a comprehensive personally curated collection can be traced back at least to Vannevar Bush [1]. Bush envisioned his “memex” as a means for storing all of a person’s “books, records, and communications.” The content comprising the memex can be subdivided into the “personal library,” books, records, and newer media created by others; and “personal archives,” the artifacts produced by the archive creator [2]. Properly curated, these collections can allow “users [to] create a personalised subset of the *information world* that they can use

when faced with information needs” [3]. New online services have fulfilled part of Bush’s dream.

Today’s personal library may not be stored locally (as Bush had proposed), but the artifacts that comprise one are nearly universally accessible anywhere with a web connection. Services (e.g., Delicious) permit easy curation of those artifacts, in many ways surpassing Bush’s vision. In contrast, a personal archive of one’s own artifacts is becoming more, not less, difficult to aggregate and keep in one place. From Facebook, to blogging platforms, to Blackboard, many services encourage users to create content directly on those sites, with no mechanisms for easily storing a copy local to the user. This content frequently can be difficult or impossible to index, bookmark or otherwise curate, and only those activities only become more cumbersome or impossible over time.

New tools have begun to be created to address this issue; however, they are almost exclusively for-profit ventures with two shortcomings: they have been created as proprietary systems, and they are developed and administered by start-ups that lack a demonstrated history of stability. The majority of new services, such as Memolane or Evertale have narrowly defined the scope of what is logged using their services, actions using a mobile phone or social media, respectively, and since neither service is open source, the user who wants to expand that scope is unable to do so. Furthermore, given the notorious volatility (and low life expectancy) of many tech startups, none of these services yet represent good candidates for long-term storage of important data.

3. DEMAND

In working with focus groups of students pursuing master’s degrees in information or library science (n=55), we saw a strong interest in a tool for aggregating and archiving the artifacts they are creating on social media sites and VLEs, but that questions of trust were critical when storing meaningful digital artifacts. Notably, while students stated that they were very interested in preserving their digital artifacts, automating that preservation was seen as critical to making it actually happen; a system that would allow users to pick what to preserve and what not to was seen as something students were unlikely to use. While the desire for such a tool may be weaker in the general population, we anticipate that archives created now will have strong, if unanticipated, future value for users.

A comprehensive personal digital curation tool would give users a resource when confronted with particular information needs, but would also help create a stronger sense of personal history by capturing digital ephemera. Perhaps more importantly, it would allow users to assert ownership of digital content that otherwise lives only on corporate servers, subject to arbitrary and unpredictable retention and metadata policies.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference’10, Month 1–2, 2010, City, State, Country.

Copyright 2010 ACM 1-58113-000-0/00/0010...\$10.00.

4. LIFETIME LIBRARY

To fill this growing need, we are developing the open source Lifetime Library. The vision of the Lifetime Library is to provide trusted storage and associated services for students and alumni for their entire lives. The university's assurance that such resources will persist will serve as a trusted covenant between the institution and its constituents. Such a service will not only help students lead successful digital lives beyond the university, but can also serve as a persistent connection to the university for alumni.

The university is offering functionally unlimited storage space during our pilot, and plans to support the system indefinitely. The metadata recorded for entered files and content is flexible and can be as formal or informal as desired. With automated policy-points available in iRODS [4], the underlying storage middleware, arbitrary programmatic actions can be taken on ingested files. This could include thumbnailing of images, cross-referencing with other concurrent events and files, notification, encryption, redistribution, or summary reporting sent back to the owner at regular intervals. Synchronization tools allow for materials stored on a person's personal computer to be added to their collection automatically and digital artifacts on third-party servers to be harvested. Activities and records from disparate locations could be aggregated and pulled back into one place and organized for easy analysis at a later date.

We have developed harvesting tools for Facebook photos and Flickr images, and are developing one for the primary VLE used by faculty and students. Images, being naturally unary files were chosen as a prime candidate for initial harvesting while strategies are developed for harvesting and archiving artifacts where the atomic unit is less obvious, such as Facebook status updates and Twitter tweets.

Harvesting tools are intended not to be limited to those created by the Lifetime Library development team. Rather, the ultimate model will be one in which the university provides a central "bucket" into which personal digital artifacts can be deposited, and an application programming interface (API) to allow interested parties to add new "pipes" to bring in artifacts from new sources. This empowers users to harness the long-term reliability of a major university for storage, while enjoying the flexibility conferred by an open source marketplace for new tools. Additionally, the API enables users to create their own tools for harvesting from additional sites while permitting university development resources to be focused on issues surrounding storage and higher-level API concerns.

5. EARLY USAGE

The Lifetime Library project has been piloting usage for one full semester with students and some alumni. Continued rollout will proceed as bugs continue to be fixed and new functionality is added. Early feedback expressed desire for more automated actions and better reporting from the system. The free storage solutions from private companies that offer many of the same features are slick and set expectations high for non-technical users.

Separating upload, or ingest, from the process of management of a personal collection has been one of our greatest goals. Removing the interface, and therefore, the friction associated with getting files backed up and into the Lifetime Library was the first step. Improving the utility of their collection through reporting, search, metadata, and analytics will be the bulk of the project moving forward. We expect many interesting questions

to present themselves as we tackle the richness of social data about a particular user, stored and organized in one place, under their control.

6. DISCUSSION

Beyond the suitability of a university to provide this type of long-term resource (by virtue of its long-term stability), beginning these types of collections just as users enter adulthood is especially appropriate. Although often considered solely the province of older adults, reminiscing is valued and engaged in throughout the life span, especially during transitional periods [5]. Despite this universality of reminiscing, most of the services on today's popular UGC sites are strongly present-focused, lacking good tools for revisiting past activities or states of mind.¹

While we expect younger adults will appreciate improved support for revisiting their digital artifacts, it is unlikely that they will spend substantial time curating this growing collection.² However, we anticipate that in later stages of their lives, users will have a more active relationship with these archives. By beginning development of the individual's personal archive at the beginning of adulthood, and by continuing to grow that archive in the decades that follow, users will have the tools they need at each stage of their likely interaction with their collection: casual browsing, sharing, curation, reminiscing, research, and storytelling.

7. CONCLUSION

Much of the discussion surrounding the information explosion and resulting information overload has focused on the impact on the role of the consumer. However, as we move toward a more "democratic" model, in which every person is alternately both a producer and a consumer, new tools are needed to help people manage their own creations alongside their libraries of artifacts created by others. The tools we are currently developing represent some early pieces of such a comprehensive information management system. As new functionality is added, both by our team and by others, the Lifetime Library vision may serve as a functional and sustainable model for personal digital curation.

8. REFERENCES

- [1] Bush, V. 1945. As We May Think. *The Atlantic Monthly*. 176, 1 (July 1945), 641-649.
- [2] Williams, P., John, J. L., and Rowland, I. 2009. The personal curation of digital objects: a lifecycle approach, *Aslib Proceedings*. 61, 4340-363.
- [3] Bruce, H., Jones, W., and Dumais, S. Information behaviour that keeps found things found. 2004. *Information Research*, 10, 1. Retrieved from October 29, 2011 from <http://informationr.net/ir/10-1/paper207.html>
- [4] Rajasekar, A., Moore, R., Hou, C., Lee, C. A., Marciano, R., de Torcy, A., Wan, M., Schroeder, W., Chen, S., Gilbert, L., Tooby, P., and Zhu, B. iRODS primer:

¹ Facebook's recently released Timeline feature is a notable exception to this rule.

² We might consider active curation to be more analogous to scrapbooking, an activity significantly more popular among older adults than younger. [6]

integrated rule-oriented data system. *Synthesis Lectures on Information Concepts, Retrieval, and Services*. 2, 1, 1-143.

[5] Parker, R. G. 1999. Reminiscence as continuity: comparison of young and older adults. *Journal of Clinical Geropsychology*, 5, 2, 147-157.

[6] Fram, E. H. 2005. The booming scrapbooking market in the USA: despite phenomenal growth, the future's unclear. *International Journal of Retail & Distribution Management*, 33, 3, 215-225.