

---

# Compound Fragmentation of Longitudinal Personal Information

**Benjamin V. Hanrahan**  
Pennsylvania State University  
bvh10@ist.psu.edu

**Manuel A. Pérez-Quiñones**  
University of North Carolina  
Charlotte  
Perez.Quinones@uncc.edu

**John M. Carroll**  
Pennsylvania State University  
jcarroll@ist.psu.edu

---

Paste the appropriate copyright statement here. ACM now supports three different copyright statements:

- ACM copyright: ACM holds the copyright on the work. This is the historical approach.
- License: The author(s) retain copyright, but ACM receives an exclusive publication license.
- Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single spaced in a sans-serif 7 point font.

Every submission will be assigned their own unique DOI string to be included here.

## Abstract

Over time, we as computer users move not only from tool to tool for a type of personal information (e.g. different email applications), but we also move through different categories of information (e.g. no one kept track of tweets before Twitter or some of us have abandoned social networking sites). This creates fragmentation not just in the moment, but over time. For example, a user might have used several different email applications or used a dozen social media applications. In this paper we extend some of our ideas for dealing with fragmentation in the moment [1] to outline a possible method for users to manage fragmentation over a longer period of time.

## Author Keywords

Personal Information Management, Fragmentation

## ACM Classification Keywords

H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous; See [<http://acm.org/about/class/1998/>]: for full list of ACM classifiers. This section is required.

## Introduction

Over time, we as computer users move not only from tool to tool for a type of personal information (e.g. different email applications), but we also move through different categories of information (e.g. no one kept track of tweets before Twit-

ter or some of us have abandoned social networking sites). This creates fragmentation not just in the moment, but over time. For example, a user might have used several different email applications or used a dozen social media applications. In this paper we extend some of our ideas for dealing with fragmentation in the moment [1] to outline a possible method for users to manage fragmentation over a longer period of time.

There are currently many silos of information in PIM, these silos lack a unifying structure and proliferate information fragmentation, a problem previously identified by the Personal Information Management (PIM) community [9]. The PIM community has documented the undesirable consequences of information fragmentation [3, 4]. At this moment fragmentation is pervasive between cloud based applications, a Dropbox folder, Gmail label, and an Evernote notebook, where there is only an informal semantic link. In our previous work we created the *Contextinator* tool [1] that leveraged ‘information views’ to address the problem of information fragmentation for cloud based applications. In this paper we leverage this approach to probe how one might do PIM over a long period of time.

Our main navigation metaphor was a project and the primary way we unified information was through leveraging structured urls. If a project utilized multiple online services (e.g., Gmail, Evernote, Dropbox), the structured urls would be saved and resuming work was easier. Using these urls pairing Evernote notebooks, Dropbox folders, and Gmail labels is not only possible, but relatively easy. Including new services was easy, as long as the product used a reasonable URL scheme.

In thinking about adapting this work to a more longitudinal use there are a few issues that we are likely to encounter. First, given the numerous services that a user moves through

in their lifetime of use and their propensity to either be taken off-line (in the case of cloud-based tools) or fall out of support (in the case of desktop tools), an exporting/archiving feature will be necessary. Second, how to index this information in a usable, memorable way must be developed perhaps being centered more on places or people, instead of project in our case. Lastly, given the ephemeral nature of many pieces of Personal Information (PI), we will need to look at which artifacts are more likely to be re-used actively by a user and which artifacts we can derive some value from for the user perhaps some of which might be social. A question this paper poses is how to pull the context out of our information?

### **Fragmentation Over Time**

Information fragmentation occurs when personal information is scattered over different devices, storage systems, or online tools and is a ‘pervasive problem in personal information management’ [9]. The siloing of information by applications is not a new problem, and it is likely to stay a problem in the foreseeable future. Users come up with different methods for mitigating these fragments: using folder hierarchies to organize related documents [8]; using a folders email to hold related messages [5]; or using virtual spaces to more physically separate projects [6].

A drawback of offloading the organization of projects on users, is that they end up maintaining duplicate, unwieldy organizational hierarchies between tools. A problem which is only compounded by time. Boardman et al. [3, 4] found that when data was compartmentalized between distinct tools, users had difficulty coordinating across different tools, and there were inconsistencies between equivalent functionality. Bergman et al.[2] framed the problem as project fragmentation, where information was fragmented into different collections without relation to the common activity

uniting them. Their solution was to use a single hierarchy to store all files of different formats under the same folder. Similarly, Jones et al. [7] suggested the development of a common structure that could be shared and manipulated by any number of tools.

However, we propose that as time passes and the number of tools ebbs and flows, fragmentation will be an even more complex problem as some services will be inaccessible or formats will have fallen out of use.

### Mitigating Fragmentation of Personal Information

Our approach to building integrated information views allowed us to address the information fragmentation that occurs on the web in the moment. A tool like *Contextinator* has the potential to create (or re-create) the context lost amid information fragmentation in today's web-based tools. But how do we support fragmentation over time?

The first problem mentioned above, export/archiving features, might be approached several different ways. For example, when a cloud service is going offline, if they are not providing an official export feature, a screen scraper could be built by users to grab this information (tooling may be necessary to avoid users having to write code). The second problem may be solved by providing more general mechanisms for recall, using people, places, and a timeline for navigation. Perhaps we could pull context out of the PI to reconstruct these navigational mechanisms and share them across collaborators. The last problem, finding value in ephemeral PI may be more difficult. Can the contexts that we pull from the PI provide some value in constructing a profile of skills and interests? How many facets of our activity (professional or otherwise) are reflected in our PI? These questions would require more research to answer, exploring what we can glean from a persons PI and how they feel it

might reflect them. A clue may be in one of the failings of our tool, several users wanted to be able to share context information with other users. It seems that information fragmentation is not just across information silos and devices, but also across collaborators and work settings. Information fragmentation may be further compounded by time as well.

### References

- [1] Ankit Ahuja, Ben Hanrahan, and Manuel A Pérez-Quiñones. 2014. Contextinator: Project-based Management of Personal Information on the Web. *Proc. 10th WEBIST* (2014), 14–23.
- [2] Ofer Bergman, Ruth Beyth-Marom, and Rafi Nachmias. 2006. The project fragmentation problem in personal information management. In *Proc. CHI '06*. ACM, New York, NY, USA, 271–274. DOI : <http://dx.doi.org/10.1145/1124772.1124813>
- [3] Richard Boardman and M. Angela Sasse. 2004. "Stuff goes into the computer and doesn't come out": a cross-tool study of personal information management. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, New York, NY, USA, 583–590.
- [4] Richard Boardman, Robert Spence, and M Angela Sasse. 2003. Too many hierarchies? The daily struggle for control of the workspace. In *Proc. HCI International 2003*, Vol. 1. 616–620.
- [5] Victor M. González and Gloria Mark. 2004. "Constant, constant, multi-tasking craziness": managing multiple working spheres. In *Proc. CHI '04*. ACM, New York, NY, USA, 113–120. DOI : <http://dx.doi.org/10.1145/985692.985707>
- [6] D. Austin Henderson, Jr. and Stuart Card. 1986. Rooms: the use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface. *ACM Trans. Graph.* 5, 3 (July 1986),

- 211–243. DOI : <http://dx.doi.org/10.1145/24054.24056>
- [7] William Jones. 2011. XooML: XML in support of many tools working on a single organization of personal information. In *Proc. iConference 2011*. ACM, New York, NY, USA, 478–488. DOI : <http://dx.doi.org/10.1145/1940761.1940827>
- [8] William Jones, Ammy Jiranida Phuwartnurak, Rajdeep Gill, and Harry Bruce. 2005. Don't take my folders away!: organizing personal information to get things done. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*. ACM, New York, NY, USA, 1505–1508.
- [9] David R. Karger and William Jones. 2006. Data unification in personal information management. *Commun. ACM* 49, 1 (2006), 77–82.