"I hereby leave my email to...": Data Usage Control and the Digital Estate

Position Paper

Stephan Micklitz*, Martin Ortlieb†, Jessica Staddon‡
* Google, Munich, Germany
micklitz@google.com
† Google, Zurich, Switzerland
mortlieb@google.com
‡ Google, Mountain View, CA, USA
staddon@google.com

***Abstract*—In most data control scenarios there is the opportunity for oversight by those who, while perhaps not directly involved in the creation of the data, understand the intended usage of the data. We argue that due to the proliferation of online data and our aging population, data owners will increasingly face requests for data access and usage when such oversight is not present because the original data owner/creator is unavailable (e.g. because of death or incapacitation). We outline the technical and user experience challenges in supporting this data usage scenario, focusing on the online service setting, and highlight some emerging research problems.**

I. Introduction

Data usage control comes with a host of challenges including policy support, verifiability and usability of the implementing systems. In most cases, however, it is assumed that the correct functionality is known, that is, for many contexts it is understood what should happen to the data, and the challenge is primarily ensuring that this actually happens. When a user dies or is otherwise unexpectedly unavailable to manage their data, even this clarity is quickly lost. To prepare for such a situation, the user would ideally have a clear data control plan that can be invoked.

Such plans are increasingly needed due to the explosion of online content. According to recent research reports from Pew [17], 91% of adults use email, 69% use an online social networking site, 22% contribute to online forum, 15% contribute to web pages or blogs, 61% bank online, 11% trade stocks online, 8% participate in an online dating site, 21% use technology to track health information [9], 65% make travel reservations online, 36% play games online and 71% buy products online. Since many of these activities involve the use of online accounts, many users have substantial digital "estates" for which they may want to plan.

From the perspective of even a single online service, the need for digital data planning can be substantial, particularly as the population ages. The world annual death rate has been close to 8 people out of 1000 in recent years [5], [13]. As an example, for GMail, which reportedly has 425M users (specifically, monthly active users [15], [18]), this rate yields a rough estimate of more than $9,300$ deaths per day, which of course, is an underestimate for Google as a whole since only a subset of Google users are GMail users. Similar estimates for Facebook have ranged as high or higher (see, for example [L12, K12]). Clearly, these are very coarse estimates as they ignore other factors that impact the actual death rate, such as the age distribution of the users of the service. That said, it seems quite likely that with the increasing popularity of online services such as these, the number of requests service providers receive for the data of demised of otherwise unavailable users will be substantial. The situation is aggravated by the fact that digital data planning is still relatively rare [19].

While death is perhaps the most common cause of this situation (e.g., see [6]) other unintended absences like extended illness or incarceration/capture should also result in execution of the user's plan. In addition, the user's plan may vary according to the circumstances in which it is executed as well as according to the specific data types it covers. For example, a human rights activist who is incarcerated may prefer a "freeze" on their account for a period of time, rather than opening access to a beneficiary immediately, and many users may wish to plan for their social networking data, blogs, email and online photo repositories quite differently.

For an online service provider to effectively support users when planning for control and usage of their data when they are no longer able to manage it, a host of technical and UX challenges must be met. The goal of this paper is to outline both the core challenges, those we see as essential to a working solution, and the challenges that will become increasingly important over time.

II. Technical and UX Challenges

Technological support for control over digital estates is a nascent area and as such users will come to any tool with some apprehension and a mix of expectations. To gain traction, it is key that tools build trust through providing reliable and secure services that meet the predominant user needs.

DOI 10.1109/SPW.2013.28

Just as account recovery tools have become a common attack vector (e.g. through password reset challenge questions [11]) a digital estate management tool is a potential vehicle for gaining unauthorized account access. To manage this risk, tool designers should consider allowing for constrained access to a users account, e.g. "read-only" access, to prevent user impersonation. In addition, delays and warnings can be used to ensure access to trustees is not given prematurely. Finally, when access to trustees is granted, strong authentication is beneficial.

One potential approach to digital data planning is simply to share passwords with trusted associates. When the account owner is no longer available, associates still have access. This is a poor security practice, though, and may be a legal or terms-of-service violation. In addition, such a practice has the potential to delay the correct triggering of digital estate management tools since account access by users who are not the primary account owner can cause the account to appear active despite a lack of use by the primary owner [6]. In addition, the fact that password sharing is often a violation of an website's terms of service, makes it a risky mechanism for digital estate planning.

For some forms of digital data, a common user need is the ability to grant read access to trustees for some period of time. Some users may prefer that their account stays readable indefinitely, others may desire that the account be deleted after some review. That said, depending on the category of data, the predominant user need may vary. For example, a user is likely to wish to provide indefinite download access to online photo albums, and potentially, write access to the same albums; whereas they are unlikely to desire the same degree of interaction with their email account. The online service provider needs to support the predominant access preferences for their particular data categories. The complaints companies like Yahoo! and Facebook have received over their treatment of deceased user accounts (see, for example, [10], [16]) attests to the difficulty of understanding user preferences.

III. Emerging Research Challenges

As technological support for digital data planning becomes more adopted, use of the tools will likely surface more subtle challenges. In particular, we highlight a few needs we anticipate will emerge.

Identifying sensitive content: Bob may specify Alice should have access to his email, forgetting the handful of emails in which he complains about Alice and/or describes intervals in their relationship that will be painful for Alice to remember. Text mining, sentiment analysis and anomaly detection may be able to flag some of these examples for Bob.

Predicting user preference for new uses: As the world changes, data use cases arise that were difficult to even imagine previously. Such use cases are unlikely to be reflected in a users stated plan for their data, yet may be cases that the user would have wanted to support. For example, a user who shares health data with an online support group, may be quite willing to have their data used for research on new diseases, assuming the data are not directly linkable to them. Can a system be trained to potentially recognize data use cases that are likely acceptable to the user, and if so, should such a system be used?

Supporting variable access to content: Some digital content, like email, is inherently personal, but other content, like online profiles and blogs, is intended for a broader audience. Some users may wish to ensure the accessibility of certain content categories, or even for certain artifacts within a category, after they lose access to their content. For example, a user may be comfortable with making their online social network account into a living memorial where friends can leave comments, but only want read access for their email account. Can support for such complex access rules be implemented in a usable fashion? Also are there any underlying rules that should govern access given the certain content may be valuable to many people [12]?

IV. Conclusion

We've argued that digital estate planning is a considerable and growing problem for online service providers. In the near-term, supporting this data control need requires addressing the abuse vectors that already exist for account recovery and understanding the main use cases for particular categories of data. As users become more comfortable with digital estate planning we anticipate a demand for services with more features and in particular, the ability to predict the sensitivity and correct data control plan for individual data artifacts.

Acknowledgment

The authors would like to thank Jess Hemerly and Eric Sachs for useful feedback on earlier drafts of this position paper.

References

[1] Number of active users at Facebook over the years. Associated Press, October 23, 2012.

[2] Brubaker, J. R., Hayes, G. R., and Dourish, J. P. (accepted). Beyond the Grave: Facebook as a site for the expansion of death and mourning. The Information Society.

[3] Brubaker, J. R., Kivran-Swaine, F., Taber, L., and Hayes, G. R. (2012). Grief-Stricken in a Crowd: The language of bereavement and distress in social media. Proc. ICWSM-12. Dublin, Ireland. June 4-8, 2012.

[4] Brubaker, J. R., and G. R. Hayes. 2011. "We will never forget you [online]": An empirical investigation of post-mortem MySpace comments. In Proceedings of the ACM 2011 conference on Computer supported cooperative work (CSCW '11), 123- 132.

[5] CIA Factbook. https://www.cia.gov/library/publications/the-world-factbook/geos/xx.html

[6] G. Fowler. Life and death online: who controls a digital legacy? The Wall Street Journal. January 5, 2013. http://online.wsj.com/article/SB10001424127887324677204578188220364231346.html

[7] J. Koetsier. 8000 Facebook members die every day. What happens to their profiles? July 20, 2012. http://venturebeat.com/2012/07/20/8000-facebook-members-die-every-day-what-happens-to-their-profiles/#IREpKRJ2k8rtdTaf.99

[8] N. Lustig. 2.89m Facebook Users Will Die in 2012, 580,000 in the USA http://www.nathanlustig.com/2012/06/06/2-89m-facebook-users-will-die-in-2012-580000-in-the-usa/

[9] S. Fox and M Duggan. Tracking for Health. January 28, 2013 http://www.pewinternet.org/Reports/2013/Tracking-for-Health/Summary-of-Findings.aspx

[10] J. Mazzone. Facebooks Afterlife. North Carolina Law Review. Volume 90, number 5. pp 1643-1685.

[11] Goodin, Dan (October 8, 2008). Son of state lawmaker charged with Palin email hack. The Register. Retrieved October 10, 2008.
[12] Graves, K. 2009. Graves, K. 2009. Social Networking Sites and Grief: An Exploratory Investigation of Potential Benefits. Dissertation at the University of Indiana at Pennsylvania.
[13] IndexMundi. http://www.indexmundi.com
[14] Massimi, M, and R. M. Baecker. 2010. A Death in the Family: Opportunities for Designing Technologies for the Bereaved. In Proceedings of the SIGCHI conference on Human Factors in computing systems (CHI 10), 1821-1830.
[15] Rani Molla. GMail finally beats Hotmail according to third party data. October 31, 2012. http://gigaom.com/2012/10/31/gmail-finally-beats-hotmail-according-to-third-party-data-chart/
[16] S. Olsen. Yahoo releases e-mail of deceased Marine. cnet. April 21, 2005. http://news.cnet.com/Yahoo-releases-e-mail-of-deceased-marine/2100-1038_3-5680025.html
[17] Online trend data - Adults. http://www.pewinternet.org/Static-Pages/Trend-Data-%28Adults%29/Online-Activites-Total.aspx
[18] Sundar Pichai. June 28, 2010. Chrome & Apps at Google I/O: Your web, everywhere. Official Google Blog.
[19] M. Zhang, C. Jennett, M. Malheiros and MA. Sasse, 2012. Data after death: User requirements and design challenges for SNSs and email providers. Proceedings Memento Mori: Technology Design for the End of Life (CHI Workshop 2012).