Failures in Sharing Personal Data on Social Networking Sites

Daniel A. Epstein

Computer Science & Engineering

DUB Group

University of Washington

depstein@cs.washington.edu

James Fogarty

Computer Science & Engineering

DUB Group

University of Washington

jfogarty@cs.washington.edu

Sean A. Munson

Human Centered Design & Engineering

DUB Group

University of Washington

smunson@uw.edu

Abstract

Sharing personal informatics data to social networking sites is a common and well-studied practice in both research and commercial applications, but there have been substantial mistakes and failures within this space that offer important lessons to application developers. We discuss three common types of failures salient in our own work, other research, and popular press stories. These failures surface important open questions to the field of personal informatics.

Author Keywords

Personal Informatics, Social Networks, Accidental Posts

Introduction

Sharing data and goal achievement on social network sites is a common practice within personal informatics. This behavior can lead to positive interactions with friends and increased accountability to goals [3,5] as well as offer free advertisement for self-tracking tools to a new audience. While prior work has offered design recommendations for improving sharing features [1,2,3], little emphasis has been placed on the mistakes in designing these features and the negative consequences they may have for self-trackers.

In this workshop paper, we focus on three trends in social sharing failures we observed within our own

Figure 1. Self-trackers often find themselves explaining inaccurate reports from their apps.

research, other research, and popular press. These trends extend beyond problems of oversharing on social network sites, and are exacerbated by the wealth of data being shared in personal informatics.

We conclude with a brief discussion of what lessons can be learned from these failures and offer high-level design recommendations. These mistakes surface opportunities where further work is necessary, and we outline some of these possibilities.

Post Inaccurately Reflects Activity

Self-trackers often express concern that the data collected by personal informatics tools and devices is inaccurate, and self-trackers wish to receive credit for all of their activity [1]. This becomes especially important when sharing, as people want to present themselves in the best possible light. This presents a dilemma for application developers. Consider a hypothetical self-tracking device that serially over-reports a positive behavior when sharing, such as a home utility tracker underestimating water usage or a pedometer overcounting steps. Such theoretical technologies may be desirable but are inherently dishonest, which is of serious concern to self-trackers when they share their data [2].

In other situations, a self-tracking device may gather and share data that is unrepresentative of self-tracker's activity. As part of an ongoing analysis of tweets made with the RunKeeper physical activity tracking application, we observed 149 tweets (3.12% of our sample) reporting activities lasting exactly zero minutes and zero seconds, such as the tweet in Figure 1 [3]. These tweets do not represent the self-tracker's actual activity, and were sometimes the subject of ridicule by

followers, receiving replies such as "*Congrats! Is that a flight of steps?*" and "*hardcore!*"

We are not certain of the cause for these tweets, but we believe that many are caused by immediately stopping an activity after starting it and then using the application to edit the details of the activity. However, RunKeeper shares a tweet as soon as the activity is recorded, and not after it is edited. Thus, the inaccurate record is shared publicly.

Post Portrays Sharer in a Negative Light

Sharing self-tracked data can often portray the sharer negatively, such as pointing out inactivity [1] or publicizing potentially embarrassing tastes in music [8]. The physical activity example may serve as an accountability feature and motivate some to be more active, but negative experiences with sharing data may also lead self-trackers to disable sharing features or even stop using a tool altogether.

Automatically sharing data can reduce barriers to sharing but also increases the chance of negative outcomes because it does not give people an opportunity to curate or filter data before it is posted publicly. Sharing too often can make one's audience "wince", even if the content is positive [6]. In systems which allow only manual sharing, such as GoalPost, self-trackers worry about posting something that appears trivial or boastful, or of "spamming their friends", and therefore often opt not to share [5].

Application developers need to consider this problem further. Perhaps the framing of these posts needs to change to avoid making the sharer appear boastful. Another possibility is to only recommend sharing major
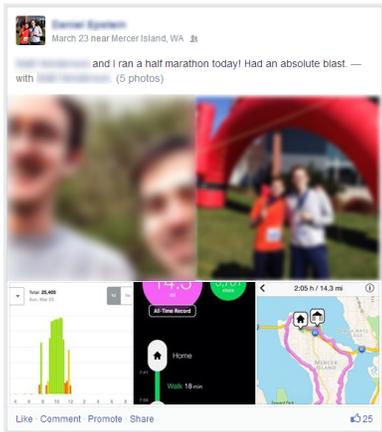
Figure 2. Tracked and shared data can give the audience more detail about major life events.
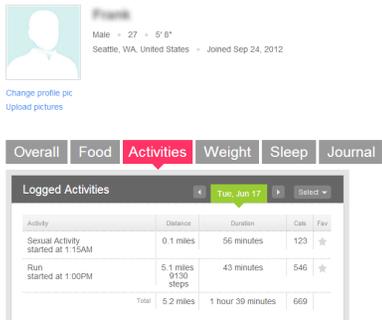

Figure 3. Logging sexual activity could have benefit for the self-tracker, but the self-tracker may not be comfortable with this data appearing in their public profile.

events (as in Figure 2), such as a personal record, a first event, or a summary of a long stretch of activity.

Post Unintentionally Exposes Data

Sharing personal informatics data could reveal patterns a self-tracker is not aware are salient in their data. To explore attitudes about privacy and sharing of physical activity data, we designed a scenario in which the self-tracker accidentally discloses a late-night trip to a local bar to a significant other through their shared activity [2]. This scenario was met with hostility, with interview participants indicating potential techniques for resolving this in the future: "*I would probably switch it to a private view*", and "*she should have left her FitBit at home!*"

The popular press has jumped on accidental disclosures like these, typically caused by personal informatics tools defaulting to making data visible publicly or to one's entire friend network. FitBit enables manual logging of any exercise including sexual activity, such as in Figure 3. This became searchable on the web and easy to tie to an individual, which caused the company to backpedal and change their default sharing settings [4]. When Facebook introduced Beacon, a feature that posted purchases to the self-tracker's Facebook wall, a purchase of an engagement ring became public before the purchaser had proposed [7].

Unshared data has value to self-trackers, who still want to maintain an accurate record of their activity. When a suggested strategy for preventing unwanted disclosures is not to track (e.g. leaving a FitBit at home, not logging sexual activity), this violates the integrity of the self-tracker's record. Sharing features should not accidentally encourage this practice.

Discussion & Conclusion

There are still many open questions in this domain, which we believe will lead to interesting conversations in a workshop environment.

Lessons Learned

Posts generated without self-tracker intervention are often the cause of problems that arise from shared personal informatics data. Requiring self-trackers to filter data prior to sharing maintains accountability for the post content and simplifies the problem of accidental disclosures. Filtering allows self-trackers to control post frequency, potentially avoiding appearing boastful, but also increases sharer burden and may cause self-trackers to not share even when they would benefit from doing so.

These failures suggest systems should not default to sharing publicly all personal informatics data. No matter how innocuous and shareable the data may seem to a tool developer, a self-tracker may use the tool in unforeseen ways. A self-tracker should not be caught in a situation where their data is accessible to others without expressly granting permission. Tools should also help self-trackers consider potential consequences of sharing, such as making public a running route that might reveal one's home residence.

Open Questions

We discuss a few unanswered questions surfaced by these sharing failures, and hope to have further discussions about them at the workshop.

Data Filtering: Filtering data prior to sharing is an effective technique for ensuring the approval of a self-tracker, but is incredibly burdensome. Imagine a

music listener selecting what to share out of thousands of songs they listened to over a year. Neither an opt-in nor opt-out approach is both effective and easy, and automated approaches will result in errors.

We also believe that it is important for application developers to aid self-trackers in realizing how others may interpret shared data. Self-trackers have enough difficulty making sense of their own data without understanding how it could be misconstrued.

Trust: Manual tracking of unverifiable data [1,4,5] and reliance on potentially unreliable autotracked data [1,8] are common and important to personal informatics. When data is presented outside of a personal, private record, questions about the accuracy and trustworthiness of this data can become important. People do not value data if they have evidence that it has been tampered with, and may distrust the sharer in the future [2]. How can sharing features enable manual tracking and data manipulation while preserving trust?

One potential solution to the problem of trust is to not share manually tracked data and to not support data manipulation, requiring self-trackers to choose whether or not to share the data exactly as it was tracked. This creates another problem: the absence of data is data. Consider a self-tracker who chooses not to share their finances for a given month after sharing for the previous year. Their sharing recipients may infer any variety of untrue and potentially negative activities about the self-tracker, such as that they exceeded their budget, they made an embarrassing purchase, or their savings have been deteriorating. Trust is in direct tension with self-tracker desire for privacy, so sharing features need to navigate this effectively.

Sharing Frequency: Sharing features are ultimately supported and promoted by commercial application developers as a viral marketing technique. Creating a large volume of posts is essential to the success of this strategy. Self-trackers and their sharing audience wish to only share important accomplishments infrequently [2,5], which is directly at odds with application developer motivations. Sharing features have yet to strike this balance.

References

1. Consolvo, S., Everitt, K., Smith, I., and Landay, J.A. Design Requirements for Technologies that Encourage Physical Activity. *CHI 2006*, 457–466.
2. Epstein, D.A., Borning, A., and Fogarty, J. Fine-Grained Sharing of Sensed Physical Activity: A Value Sensitive Approach. *UbiComp 2013*, 489–498.
3. Epstein, D.A., Jacobson, B.H., Bales, E., McDonald, D.W., and Munson, S.A. "It's just another stupid post": Design Dimensions to Improve Social Sharing in Personal Informatics. *Under Review*.
4. Hill, K. Fitbit Moves Quickly After Users' Sex Stats Exposed. *Forbes*, 2011. http://www.forbes.com/sites/kashmirhill/2011/07/05/fitbit-moves-quickly-after-users-sex-stats-exposed/.
5. Munson, S.A. and Consolvo, S. Exploring Goal-Setting, Rewards, Self-Monitoring, and Sharing to Motivate Physical Activity. *PervasiveHealth 2012*, 25–32.
6. Munson, S.A., Lauterbach, D., Newman, M., and Resnick, P. Happier Together: Integrating a Wellness Application into a Social Network Site. *Persuasive Technology*, (2010), 27–39.
7. Pepitone, J. Facebook broadcasts engagement ring purchase. *CNN Money*, 2010. http://money.cnn.com/galleries/2010/technology/1012/gallery.5_data_breaches/3.html.
8. Silfverberg, S., Liikkanen, L.A., Lampinen, A., and Hall, S. "I'll press Play, but I won't listen": Profile Work in a Music-focused Social Network Service. *CSCW 2011*, 207–216.