Disclosure and Timeliness: How much time must pass before it's okay to share?

Presented by: Terrell Russell, Jacob Kramer-Duffield

University of North Carolina at Chapel Hill

Research has repeatedly shown that computer-mediated communications (CMC) lead to higher levels of disclosure of personal information (Tidwell and Walther 2002). Recent studies have examined the role of increasingly common social media and social network services (SNS) on disclosure in a variety of contexts (Mazer et al. 2007; Tufekci 2008). [The combination of personal demographic data, taste preferences, public disclosure of friend networks and now increasing usage of tools for instantly updating status (e.g., Twitter, Facebook) has, we believe, fundamentally altered users' understanding of the temporality of information and its (semi-)permanence.] Some studies have begun to look at longer-term life-cycle considerations of data disclosure within certain contexts (Williams 2008) but there has not yet been a comprehensive effort to gauge the effect of timeliness on self-disclosure for various types of personal data, especially for explicitly time-identified disclosures. Our research will examine this issue through a survey mechanism gauging user reactions and perceptions.

The self-published data we are interested in is, first and foremost, published by people. These people can be classified in numerous ways, and for this analysis we will focus on age, gender and an index of sociability to distinguish users and characterize their behaviors in terms of disclosure. The sociability index will be derived from questions in the CPI sociability scale.

The types of data being shared that we are interested in analyzing can be clustered across two different determining factors. First, the data can be clustered with regards to temporality and rate of change. Items like proximity data, event participation, status updates tend to change at a fairly quick rate when compared to friend listings, gender, relationship status, or taste preferences. The second factor for clustering data types revolves around intimacy of the data being shared. Few would probably argue that a status update referring to a silly thing overheard on the bus is equally as disclosing as a bit of personal health data. Among the data we are interested in working with are photographs, video, friend list activity (friending), health data, proximity and location, politics, religion, and status updates.

In addition to the people doing the disclosing, and the data being disclosed, we're interested in the audience with which the disclosure is being shared. Numerous studies have shown differing levels of intimacy and disclosure based on varying closeness of audience. Whether the intended audience is close friends and family, acquaintances and familiar strangers, or everyone on the public internet, we hope to get a sense of how far a piece of disclosed information will be perceived to travel – its reach or distance.

The final and most central factor of interest in these investigations is the temporal element - namely, how does the passage of time affect acceptability of personal information disclosure? The

time scale we'll be using will range from "now/instant" to "never" and will be designed to capture a representative portrait of user behavior with regards to online disclosure.

We will conduct our survey and tests using Amazon's Mechanical Turk [http://www.mturk.com], which has been shown to return reliable results in studies of CMC and interactivity. We will be basing our survey methodology on validated scales for each category and variable.

We expect to see that, with time, data that was originally too private or too revealing will mellow and become acceptably shareable. We expect to see this trend across the different types of shared data as well as the range of intended audiences. We expect the temporal delay to have the effect of sufficiently anonymizing the data so that we hear "it's about who I was, not who I am".