

File 20110114.0131: Weekly activity report 0171:

weekly activity report 171 (loughry)

Joe Loughry

Sent: 14 January 2011 01:31

To: Niki Trigoni; Andrew Martin; Joanna Ashbourn

Cc: otaschner@aol.com; anniecruz13@gmail.com; andrea@hpwtdogmom.org;

chip.w.auten@lmco.com; edloughry@aol.com; diane@dldrncs.com;

Joe Loughry; mmcauliffesl@comcast.net; tom.a.marso@lmco.com

Weekly activity report no. 20110113.1239 (GMT-7) sequence no. 0171, noughth week HT

I received the GSS report from Dr Martin. He noted in his report that the assessors affirmed the direction and approach of my research despite the fact that they wanted to see different evidence. I am following the assessors' advice and intend to give them precisely what they asked for, subject to the additions that my supervisor and I talked about in our last meeting. Dr Martin said in the GSS report that he believes I can finish within the deadline without needing an extension, and that I will make a further application for confirmation of status in the summer along with a draft thesis requiring relatively little additional work. To that end, I have been reading up on Grounded Theory.

One of the first things I found was that the uneasiness I had with the assessors' position---vis--vis the proper use of data---is really a point of contention in the field of qualitative research. Grounded theory is a rejection of the scientific method by its proponents who claim that the resulting theories, being 'grounded' in data collected without regard to a prior hypothesis, are thereby more defensible than what Glaser calls the 'unverifiable' hypotheses of quantitative methods [Glaser and Strauss, pp. 15--18]. Of course this turns Popperian falsifiability on its head and explains a lot of the tension I felt in that viva. To me it feels like doing science willfully wrong. Dr Martin and I have talked about it in the past. My assessors have a very different view of what qualifies as proof, he said, but that reflects the different camps that exist in computer science and even within the department. I can see, however, the attraction of grounded theory for this particular problem. The assessors, as I said before, convinced me to use it and made it plain that they want the data I have collected. The methodology is a fascinating new thing to study.

I began with 'The Discovery of Grounded Theory' by B.G. Glaser and A.L. Strauss (New Brunswick: Aldine, 1967) which does a good job of explaining why qualitative methods ought to be used. It's an old book but valuable because it isn't crammed full of fiddly little details about different software packages. I plan to get into the coding manuals by Charmaz (2006) and Saldaa (2009) as the next thing. No results to report yet, as I am still learning the methodology.

In Security Reading Group this week, John Lyle introduced the paper 'Personal privacy through understanding and action: five pitfalls for designers' by Lederer et al. [J. Pers. Ubiqu. Comput. 8(6), pp. 440--454, 2004]. John, myself and Dr Martin were present; Cornelius arrived a little later. The authors define five pitfalls of privacy settings with regard to the user interface. The pitfalls are: (1) obscuring potential information flow, (2) obscuring actual information flow, (3) emphasising configuration over action, (4) lacking coarse-grained control, and (5) inhibiting existing practice. It is a well-written paper with relevance to Cross Domain Systems in addition to the obvious applicability to social networks, which did not exist in their present form when this paper was written. John told us about the paper in the context of Webinos

and his efforts to avoid the five pitfalls in the design of that system. The paper has been well cited, although unaccountably it is not widely known. It is old enough that it pre-dates Facebook and Twitter but it anticipated the issues with location tracking, privacy configuration, and even Wikileaks, I believe.

The paper begins with a philosophical but ultimately empty definition of privacy. I thought the introduction was wonderful but then the paper got a little lost for a couple of pages before returning to the good stuff. The citation to Westin (1967), I thought, was great as the paper traces the development of privacy policy (which affects both European and American law) all the way back to the 'fair information practices'---something missed in many papers on the subject of privacy. John began by linking the five pitfalls to the design of Webinos and asking the question why the authors chose to present their findings as pitfalls instead of positive principles. This prompted a discussion of positive vs negative examples, the value of negative examples, and John's goal to make sure that Webinos does not suffer from the same pitfalls. I suggested turning around three of the pitfalls into a design principle that (1) makes privacy-relevant information flow open and transparent by default, (2) provides immediate feedback to the user about information flows (perhaps with symbology), and (3) solicits the user to continuously fine-tune his privacy settings. It would look cool to present the information flows using MIL-STD-2525B or NTDS symbology, i.e., circle for friendly, square for unknown, diamond for hostile, cross for neutral; upper half only for airborne, lower half only for subsurface targets.

Plausible deniability is another interesting idea in this paper. John noted that plausible deniability is an underrated feature of mobile phones, and that Twitter respects the principle. Twitter has the potential to be extremely fine-grained, but because it's not an automatic system, users can control information flow simply by choosing to do nothing. Automatic tools (for example, location reporting badges) do not respect this principle. I offered that the 'precision dial' (Section 8.2) needs another setting: to deliberately lie. This prompted a discussion of audit trails, inferring of information from lack of data, what it might tell other people when they see you making changes to your own privacy settings---possibly or apparently in response to an immediate event---and whether deleting a tweet actually communicates an implicit message. The whole Wikileaks controversy, in my opinion, is about precisely the same level-of-privacy question as apply to Twitter users. Perhaps, I suggested, as a design principle it would be useful to introduce classes of device privacy behaviour to end users in terms of 'this device behaves like a mobile' or 'this device behaves like a location tracking badge' or 'this system behaves like Twitter'.

Dr Martin noted that some people in the current decade are extremely reluctant to disclose their sort code and bank account numbers to a web site, yet a few years ago these same people scattered cheques around like confetti containing the same information. John suggested that people are more aware now of the potential for misuse. In one of the authors' examples---adjusting the precision of reported information---this is just the sort of thing that Cross Domain Systems do in some cases, another connection to my research. Dr Martin brought up the collection of personal information including purchasing habits and search history by on-line retailers for the purpose of marketing and targeted advertising, and wondered whether in future people might attempt to turn this around by gaming the system in order to obtain advantage over the companies. We next discussed Amazon.com recommendations and the changes in behaviour that occur when people are aware of surveillance, such as in the presence of visible CCTV cameras. Amazon.com is an interesting case because their

search history remembers \emph{everything}; I do sometimes deliberately not search for uncharacteristic things on Amazon because I know it will immediately be incorporated into recommendations, skewing the results. Dr Martin offered some specific examples of his own.

I wouldn't mind having a monitoring process that would watch my activity and infer, minute-by-minute, what project I am working on and assign that time to a notional billing system for time tracking purposes. Dr Martin noted that the necessary information might already exist in (or at least be accessible to) the window manager and the operating system. John observed that people are known to modify their behaviour when they are aware that they are under surveillance; this prompted a discussion of timekeeping and monitoring in the workplace by employers. I remember a very early paper from PARC describing their experience with RF location tracking badges; they intended it for use routing telephone calls automatically to the office where a person was, but discovered almost immediately that users took off their badges and left them in the toilet. This was recognised as privacy awareness leading to implicit modification of behaviour back in the 1970s.

John defined privacy as the right to be left alone. Dr Martin recalled a paper (I have no reference for this yet) cataloguing about as many definitions of privacy as Dieter Gollman had for trust in his paper in 2005. Dr Martin's definition of privacy echoed European laws in that he called it 'ensuring that the subject has control over how the information is used'. That is the fundamental difference between European and American privacy laws, in my opinion: this notion of who controls. In the USA, control or ownership vests in responsibility for protecting PII because it was felt that distributing ownership in a fine-grained fashion would be prohibitively expensive given the size of the population. An awful lot of people, said Dr Martin, conflate privacy with secrecy.

Next week, Shamal has offered a paper on Mis-usability Cases. Cornelius is updating the wiki and planning to migrate it to the Comlab wiki with SSO access soon.

Dr Martin is running a seminar series beginning Monday with an interesting set of guest speakers next term. I will attend by teleconference.

I did some work for the Oxford University Scientific Society this week, setting up their termcard on the web and helping with the Hilary schedule. The Air Force sponsor emailed earlier in the week asking for an emergency report containing status and a metrics baseline for the probabilistic redaction project; I got it back to him before his deadline.

My current tasks, in priority order, are:

1. Reading a stack of books from Amazon on qualitative data analysis, grounded theory and two coding manuals, plus a survey article.
2. Define the set of codes and categories for R'', keeping in mind that I want to capture discrete events as well.
3. Figure out if I can use the chronological record in my lab notebook as a source for the 'memo writing' activity that occurs later.
4. Try using the Tasks feature of Google Calendar as a project planning tool. I found an article purporting to show how to do it.
5. Outline the survey journal article that the assessors asked for. I still need to find some good examples of how to write a survey article.

Joe Loughry
Doctoral student in the Computing Laboratory,
St Cross College, Oxford

End of WAR 0171.

References