KEEPING SOFTWARE ENGINEERS ACCOUNTABLE

THE STRUGGLE FOR DATA PRIVACY IN THE U.S.

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Jason Tufano

November 2, 2020

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS
Peter Norton, Department of Engineering and Society
Aaron Bloomfield, Department of Computer Science

General research problem

How can personal data be kept private in the increasingly digital world?

Software engineers develop the systems that tech companies use to protect the vast user data they store. Despite this responsibility, some users find end products to be lackluster at best. In 2017, only 20 percent of banks in the United Kingdom were "highly confident in [their] ability to detect a data breach" (McCallion, 2017). Yet users have no other option but to hope that these systems protect their data. In a study of 474 data breaches between March and June 2020, Verizon found that while 45 percent of the attacks featured hacking, 22 percent of the attacks featured a social element, such as phishing (Verizon, 2020). Thus, software engineers, data collection companies, and software users are all responsible for keeping personal data secure.

Keeping software engineers accountable: A computer science ethics course

How can software engineers create software that protects the rights and data of its users?

This will be a Computer Science capstone project with Aaron Bloomfield as the temporary technical advisor. The technical advisor for the project is likely to change in the spring 2021 semester, when the project will be completed. I expect there to be no collaborators on the project.

Software engineers are responsible for developing many of the systems we rely on every day; therefore, it is reasonable to expect them to be cognizant of all of the impacts of their work and to consider these impacts during design and implementation of software. However, discussion of ethics in computer science courses often only scratches the surface of the many issues at play for software engineers. Computer science education researchers disagree with this

approach, stating that "teaching computer scientists to identify and address ethical problems starting from the design phase" is as important as teaching them how to develop algorithms (Grosz et al., 2019). For many computer science students, the current best resource for ethical software development is the ACM Code of Ethics and Professional Conduct, which is "designed to inspire and guide the ethical conduct of all computing professionals" (ACM, 2018). While the ACM Code of Ethics provides a good starting point for engineers aspiring to ethical software development, critics say that the ACM's framing of ethical problems as a "matter of personal choice and responsibility … obscures more complex interactions between ethics and technology", and thus would not be sufficient as the sole ethical resource for computer scientists (Vakil & Higgs, 2019).

The main objective of this project is to establish a workable starting point for the creation of a computer science ethics course. As the solution must be in the form of a modification to an existing UVA computer science course or a new UVA computer science, the goal is to create a syllabus for a computer ethics class, along with approximately one week of course material. To this end, current UVA computer science curriculum will be reviewed alongside ethics curriculum (with emphasis on computer science ethics courses) to determine the best way to incorporate an ethics course into the UVA computer science programs. If the project is successful, the material produced will be a resource for the computer science department to consider. Ideally, the department will recognize the importance of teaching ethics to computer scientists and will incorporate the results of my project into the department's offerings in some form.

The decline of privacy and the rise of data collection

How do data brokers defend their reputation despite their controversial business model?

Although most people say they value their online privacy, few Facebook users quit the platform following the Facebook-Cambridge Analytica scandal in 2018 (Hinds et al., 2020). Google also collects vast user data, yet few Google users resort to other search engines that better protect their privacy. For social media users, network effects limit the competitive opportunities for more secure alternatives to Facebook (Belvaux, 2011), and online social engagement may be more important to them than privacy protection (Zhu & Chen, 2015). Nevertheless, neither Facebook nor Google can afford to ignore objections to their data collection practices. Both companies must balance the attractions of unlimited data monetization against the risks of demands for public regulations. This means the companies must strive to represent themselves as responsible corporate citizens, but without curtailing profitable practices more than necessary.

This is not the first time companies have needed to defend their public image. Villas-Boas (2004) studied the effect of price cuts on consumer loyalty, and found that some consumers value knowledge of a brand over changes in price, which may explain why Facebook has remained popular despite the rise of competitors and the occurrence of data scandals (Hinds et al., 2020). Zhu and Chen (2015) found that targeted advertisements that tie in to the human needs that social media fulfills are more effective than those that do not.

Facebook, Inc., is a major data collector. CEO Mark Zuckerberg has admitted that Facebook does not "have a strong reputation for building privacy protective services," but claims that is has "repeatedly shown that we can evolve to build the services that people really want" (Zuckerberg, 2019). Many platform users are skeptical of such claims, though most continue to

use it. Some users limit what they post, such as sales director David Garvey who said "right off the bat I ... understood that anything you post on there can be seen pretty much by anyone at any time." Other users, such as 98-year-old Shirley O'Key, say data collection is expected: "nothing is private. How stupid can we be?" (Beck, 2018).

The Electronic Privacy Information Center (EPIC) is a data privacy advocacy that promotes awareness of data privacy and urges Congress to pass laws that better protect user data. In 2019, EPIC's president and policy director called on Congress to establish a U.S. Data Protection Agency (Rotenberg & Fitzgerald, 2019). EPIC also supports laws that limit data collection options of corporations, such as the Telephone Consumer Protection Act (TCPA), arguing that it "invad[es] the privacy of American homes" (Barr v. AAPC, 2020). The Electronic Frontier Foundation (EFF) provides tools that reveal what groups track users' data, and to what extent. For example, EFF offers a "Who Has Your Face?" quiz. When users complete the quiz, a message urges them to join the fight to "ban facial recognition" (EFF, 2020).

The Insights Association is a trade association that lobbies legislators to protect the data analytics industry. The group opposes broad interpretation of the TCPA, claiming it "threatens to impose liability on the businesses for communications that are helpful to, and desired by, consumers" (Fienberg, 2020). The Internet Association (IA), founded by Google, Amazon, and Facebook, also represents data collectors' interests. IA claims to defend user privacy, but it was critical of California's 2018 Consumer Privacy Act, demanding that "policymakers work to correct the inevitable, negative policy and compliance ramifications [of] this last-minute deal" (IA, 2018).

References

ACM (2018). Association for Computing Machinery. ACM Code of Ethics and Professional Conduct. https://www.acm.org/code-of-ethics

Beck, J. (2018, June 08). People Are Changing the Way They Use Social Media. *Atlantic.* https://www.theatlantic.com/technology/archive/2018/06/did-cambridge-analytica-actually-change-facebook-users-behavior/562154/

Belvaux, B. (2011, March). The Development of Social Media: Proposal for a Diffusion Model Incorporating Network Externalities in a Competitive Environment. *Recherche Et Applications En Marketing (English Version)*. ProQuest.

Barr v. AAPC (2020). Brief of Amici Curiae Electronic Privacy Information Center et al. in Support of Petitioners, Barr v. American Association of Political Consultants, Inc., United States Supreme Court No. 19-631.

EFF (2020, March 19). Electronic Frontier Foundation. Are Your Identification Photos in a Face Recognition Database? https://www.eff.org/press/releases/are-your-identification-photos-face-recognition-database

Fienberg, H. (2020, September 14). Insights Association Joins TCPA Challenge at the Supreme Court: Insights Association. https://www.insightsassociation.org/article/insights-association-joins-tcpa-challenge-supreme-court

Grosz, B. J., Grant, D. G., Vredenburgh, K., Behrends, J., Hu, L., Simmons, A., & Waldo, J. (2019). *Embedded EthiCS*. EBSCOhost.

Hinds, J., Williams, E. J., & Joinson, A. N. (2020). "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies, 143*, 102498. ScienceDirect.

IA (2018, June 28). Internet Association. Statement on the Enactment of California Privacy Legislation. https://internetassociation.org/news/statement-enactment-california-privacy-legislation/

McCallion, J. (2017, February 2). Data breach news: Most UK banks don't believe they can detect a data breach. *ITPro.* ProQuest.

Rotenberg, M., & Fitzgerald, C. (2019, December 3). Legislative Proposals to Protect Privacy [PDF]. Washington, DC: Electronic Privacy Information Center.

Vakil, S., & Higgs, J. (2019, March). It's About Power: A call to rethink ethics and equity in computing education. *Computers & Applied Sciences Complete.* EBSCOhost.

Verizon. (2020). Data Breach Investigations Report. https://enterprise.verizon.com/resources/reports/dbir/

Villas-Boas, J. M. (2004). Consumer Learning, Brand Loyalty, and Competition. *Marketing Science, 23*(1), 134-145. JSTOR.

Zhu, Y., & Chen, H. (2015). Social media and human need satisfaction: Implications for social media marketing. *Business Horizons, 58*(3), 335-345. ScienceDirect.

Zuckerberg, M. (2019, March 06). A Privacy-Focused Vision for Social Networking. https://www.facebook.com/notes/mark-zuckerberg/a-privacy-focused-vision-for-social-networking/10156700570096634/