EXHIBIT A

IN THE UNITED STATES DISTRICT COURT

FOR THE NORTHERN DISTRICT OF CALIFORNIA

IN RE LINKEDIN USER PRIVACY LITIGATION

Case No. 12-cv-03088-EJD

DECLARATION OF DR. SERGE EGELMAN IN SUPPORT OF PLAINTIFF'S SECOND AMENDED CONSOLIDATED CLASS ACTION COMPLAINT

[Hon. Edward J. Davila]

DECLARATION OF DR. SERGE EGELMAN

Pursuant to 28 U.S.C. § 1746, I, Dr. Serge Egelman, hereby declare and state as follows:

1. I am a resident of the State of California, and am a computer scientist. A copy of my curriculum vitae is attached hereto as Exhibit 1. I am over the age of eighteen and am fully competent to make this declaration. This declaration is based upon my personal knowledge, except where expressly noted otherwise. If called upon to testify as to the matters averred herein, I could and would competently do so.

Initial Bases for Analysis

2. For purposes of this declaration, I have been asked to analyze users' data security and privacy expectations when using LinkedIn's Premium Subscriptions, the role data security and privacy practices play in consumer purchasing decisions, as well as LinkedIn's actual data security and privacy practices and reactions to those practices.

3. I have reviewed a significant number of documents in this case including the Amended Consolidated Complaint, (Dkt. 54), LinkedIn's Motion to Dismiss, (Dkt. 59), and the Court's March 5, 2013 Dismissal Order, (Dkt. 70). I have also reviewed the existing academic literature concerning behavioral economics and privacy and security practices, as well as the underlying data from my own previous studies of privacy and security practices. I also reviewed popular media coverage of the LinkedIn data breach, and I conducted two online surveys in order to determine consumer attitudes toward social networks and data security practices.

Analysis

4. By reviewing the aforementioned materials, I was able to reach several conclusions. First, through a review of the existing academic literature, I determined that consumers incorporate data security and privacy concerns, costs, and benefits into their purchasing and consumption decisions, and that consumers are often willing to pay a premium for information security.

5. Second, through a survey I conducted the week of April 1, 2013, I determined that when consumers pay for a "premium" social networking service, they expect their information to be protected with a heightened level of security, and that, at a bare minimum, industry-standard security protocols will be used to guard their information.

6. Third, through a survey conducted the week of April 22, 2013, I determined that an internet service using industry-standard security practices has higher utility to consumers than a service with substandard security. I also determined that when consumers are evaluating the utility of a website or internet service, privacy and security concerns factor heavily into that evaluation, and that consumers will choose a website or internet service with industry-standard security practices over an otherwise identical service with substandard security.

7. Fourth, I consulted existing literature and research, along with news reports and LinkedIn's own admissions, and determined that LinkedIn's data security practices fall far short of industry standards, and indeed have been outdated since at least 2006.

8. Fifth, based on my own previous research, real-world examples, and media responses to the LinkedIn breach, I determined that had LinkedIn disclosed its true security practices, rather than representing that it used "industry standards," LinkedIn's Premium Subscribers and the public more generally would have learned of LinkedIn's substandard security practices, and would have considered that information as part of their purchasing decisions.

Conclusion

9. Through my investigation, I concluded that consumers factor data security practices into their purchasing decisions, that when LinkedIn's Premium Subscribers paid for their subscriptions they expected industry-standard data security, and that online services offering industry standard security are more useful and valuable to consumers than services offering substandard security. My research also showed that LinkedIn's security practices fell far below industry standards, and that had LinkedIn disclosed its true security practices, its current and potential Premium Subscribers would have learned of those disclosures and factored them into their purchasing decisions.

10. Thus, it is my conclusion that, had LinkedIn disclosed that it used outdated, unsalted SHA-1 hashing to protect users' personal information, its Premium Subscribers would have found out, and would have been less willing to purchase Premium Subscriptions.

I declare under penalty of perjury under the laws of the United States of America that, to the best of my knowledge, the foregoing is true and correct. Executed on April 30, 2013 at Paris, France.

Dr. Serge Egelman

EXHIBIT A-1

Serge Egelman

731 Soda Hall
Berkeley, CA 94720
USA

Email: serge@guanotronic.com

Education

- **PhD in Computation, Organizations, and Society**, May 2009
  School of Computer Science, Carnegie Mellon University
- **BS in Computer Engineering**, May 2004
  School of Engineering and Applied Science, University of Virginia

Refereed Journal Publications

- The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. Information Systems Research (ISR), 22(2), June 2011, pp. 254-268 (with J. Tsai, L. Cranor, and A. Acquisti). *Best Published Paper Award!*
- P3P Deployment on Websites. Electronic Commerce Research and Applications (ECRA), Autumn 2008 (with L. Cranor, S. Sheng, A. McDonald, and A. Chowdhury).
- The Real ID Act: Fixing Identity Documents with Duct Tape. I/S: A Journal of Law and Policy for the Information Society, 2(1), Winter 2006, pp. 149-183 (with L. Cranor).

Refereed Conference Papers

- The Importance of Being Earnest [in Security Warnings]. Financial Cryptography and Data Security. 2013 (with S. Schechter), to appear.
- Does My Password Go up to Eleven? The Impact of Password Meters on Password Selection . CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 (with C. Herley, A. Sotirakopoulos, I. Muslukhov, and K. Beznosov), to appear.
- My Profile Is My Password, Verify Me! The Privacy/Convenience Tradeoff of Facebook Connect . CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013, to appear.
- Android Permissions: User Attention, Comprehension, and Behavior. Proceedings of the 2012 Symposium on Usable Privacy and Security (SOUPS). July 2012 (with A. P. Felt, E. Ha, A. Haney, E. Chin, and D. Wagner). *Best Paper Award!*
- Facebook and Privacy: It's Complicated. Proceedings of the 2012 Symposium on Usable Privacy and Security (SOUPS). July 2012 (with M. Johnson and S. Bellovin).
- Oops, I Did It Again: Mitigating Repeated Access Control Errors on Facebook. CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 (with A. Oates and S. Krishnamurthi).
- Of Passwords and People: Measuring the Effect of Password-Composition Policies. CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 (with S. Komanduri, R. Shay, P. G. Kelley, M. Mazurek, L. Bauer, N. Christin, and L. F. Cranor). *Best Paper Nominee!*
- It's All About The Benjamins: An empirical study on incentivizing users to ignore security advice. Financial Cryptography and Data Security. 2011 (with N. Christin, T. Vidas, and J. Grossklags).
- Crying Wolf: An Empirical Study of SSL Warning Effectiveness. The 18th USENIX Security Symposium. 2009 (with J. Sunshine, H. Almuhimedi, N. Atri, and L. Cranor).
- It's No Secret: Measuring the reliability of authentication via 'secret' questions. The 2009 IEEE Symposium on Security and Privacy (with S. Schechter and A.J. Brush).
- It's Not What You Know, But Who You Know: A social approach to last-resort authentication. CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009 (with S. Schechter and R. Reeder).
- Timing Is Everything? The Effects of Timing and Placement of Online Privacy Indicators. CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009 (with J. Tsai, L. Cranor, and A. Acquisti).
- Family Accounts: A new paradigm for user accounts within the home environment. CSCW '08: Proceedings of the 2008 Conference on Computer Supported Cooperative Work. 2008 (with A.J. Brush and K. Inkpen).
- You've Been Warned: An Empirical Study on the Effectiveness of Web Browser Phishing Warnings. CHI '08: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 (with L. Cranor and J. Hong). *Best Paper Nominee!*
- Phinding Phish: An Evaluation of Anti-Phishing Toolbars. NDSS: Proceedings of the ISOC Symposium on Network and Distributed System Security. February 2007 (with Y. Zhang, L. Cranor, and J. Hong).
- An Analysis of P3P-Enabled Web Sites among Top-20 Search Results. Proceedings of the Eighth International Conference on Electronic Commerce. August 2006 (with L. Cranor and A. Chowdhury).
- Power Strips, Prophylactics, and Privacy, Oh My!. Proceedings of the 2006 Symposium On Usable Privacy and Security (SOUPS). July 2006 (with J. Gideon, L. Cranor, and A. Acquisti).

Refereed Workshop Papers

- I've Got 99 Problems, But Vibration Ain't One: A Survey of Smartphone Users' Concerns. The 2nd Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM). October 2012 (with A. P. Felt and D. Wagner).

- How to Ask for Permission. The 7th USENIX Workshop on Hot Topics in Security (HotSec '12). August 2012 (with A. P. Felt, M. Finifter, D. Akhawe, and D. Wagner).
- Choice Architecture and Smartphone Privacy: There's A Price for That. Workshop on the Economics of Information Security (WEIS). June 2012 (with A. P. Felt and D. Wagner).
- How Good Is Good Enough? The Sisyphean Struggle for Optimal Privacy Settings. CSCW 2012 Workshop on Reconciling Privacy with Social Media. February 2012 (with M. Johnson).
- Toward Privacy Standards Based on Empirical Studies. W3C Workshop on Web Tracking and User Privacy. April 2011 (with E. McCallister).
- Please Continue to Hold: An empirical study on user tolerance of security delays. Workshop on the Economics of Information Security (WEIS). June 2010 (with D. Molnar, N. Christin, A. Acquisti, C. Herley, and S. Krishnamurthi).
- Tell Me Lies: A Methodology for Scientifically Rigorous Security User Studies. Workshop on Studying Online Behaviour at CHI'10. April 2010 (with J. Tsai and L. F. Cranor).
- The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. Workshop on the Economics of Information Security (WEIS). June 2007 (with J. Tsai, L. Cranor, and A. Acquisti).
- Security User Studies: Methodologies and Best Practices. CHI '07 Extended Abstracts on Human Factors in Computing Systems. April 2007 (with J. King, R. Miller, N. Ragouzis, and E. Shehan).
- Studying The Impact of Privacy Information on Online Purchase Decisions. Workshop on Privacy and HCI: Methodologies for Studying Privacy Issues at CHI'06. April 2006 (with J. Tsai, L. Cranor, and A. Acquisti).

Book Chapters and Magazine Articles

- Crowdsourcing. To appear in *Ways of Knowing in HCI*, J. Olson and W. Kellogg (Eds.), to be published by Springer (with E. Chi and S. Dow).
- Helping Users Create Better Passwords. *;login:*. December 2012 (with B. Ur, P. G. Kelley, S. Komanduri, J. Lee, M. Maass, M. Mazurek, T. Passaro, R. Shay, T. Vidas, L. Bauer, N. Christin, L. F. Cranor, and J. Lopez).
- Conference Report: SOUPS 2006. *IEEE Security & Privacy*. November/December 2006 (with J. Tsai).
- Conference Report: 14th USENIX Security Symposium. *;login:*. December 2005 (with K. Butler, M. Chow, J. Duerig, B. Hicks, F. Hsu, S. Kelm, and M. Rajagopalan).
- Conference Report: 13th USENIX Security Symposium. *;login:*. December 2004 (with A. AuYoung, E. Cronin, M. Dougherty, R. Greenstadt, S. Kelm, Z. Liang, C. Mano, N. Smith, A. Raniwala, T. Whalen, and W. Xu).
- Suing Spammers for Fun and Profit. *;login:*. April 2004.
- Installation. *Peter Norton's Complete Guide to Linux*. Macmillan Computer Publishing. 1999.
- User Administration. *Peter Norton's Complete Guide to Linux*. Macmillan Computer Publishing. 1999.

Research Experience

Scientist
University of California, Berkeley
September 2011-present

I am currently working with David Wagner's research group to examine privacy and security issues on mobile devices (e.g., smartphones). Specifically, we are examining how users make decisions to install particular applications and how to better alert them to potential malware. We are in the process of creating a new architecture for prompting users when an application requests certain hardware or software abilities.

Scientist
NIST
August 2010-July 2011

I helped design and conduct studies to examine how users interact with authentication systems, specifically password and token-based systems. I co-organized a workshop on the NIST campus to discuss ways in which usable security research and techniques could be formally integrated into the development process, as well as reviewed grant proposals for NIST funding.

Postdoctoral Research Associate
Brown University
August 2009-August 2010

I worked with Shriram Krishnamurthi on creating better interfaces for policy authors to specify access control policies. We conducted studies to determine common policy errors, the causes of these errors, and the types of interfaces that policy authors currently use. We developed a new policy authoring interface that allows users of social networking websites to interactively specify policies in order to more easily detect and clarify ambiguities. We designed and conducted a usability study to validate our tool.

Research Assistant
Carnegie Mellon University
June 2004-May 2009

While pursuing a PhD under the direction of Dr. Lorrie Cranor in the Computation, Organizations, and Society program at CMU, I focused primarily on the usability of privacy and security systems. Areas that I worked in included creating more effective web browser trust indicators, creating usable privacy tools, Internet anonymity, and detection and prevention of phishing attacks. My dissertation is entitled "Trust Me: Designing Trustworthy Trust Indicators." My committee consisted of Lorrie Cranor (chair), Jim Herbsleb, Jason Hong, and Steve Bellovin (Columbia University).

Research Intern
Microsoft Research
July 2008-October 2008

During my second internship at MSR, I conducted two user studies with Stuart Schechter. We first looked at using social networks as a means for authenticating webmail users who had forgotten their passwords. We tested the usability of our

system as well as how susceptible it would be to various attacks. Additionally, I assisted the Internet Explorer team with new designs for their security warnings based on my research. We tested the new warnings in the laboratory using an eye tracker.

Research Intern
Microsoft Research
January 2008-April 2008

I was an intern at MSR working with A.J. Brush and Kori Inkpen on user account models for shared family computers. We examined why the current user account model does not work on computers shared by trusted individuals (i.e. communal home computers) and developed a more appropriate model. I implemented our prototype in C# and ran a usability study. This work was published at the 2008 Computer Supported Cooperative Work (CSCW) conference.

Research Intern
Xerox PARC
June 2006-September 2006

During the summer of 2006, I worked with Jim Thornton in the Computer Science Lab (CSL) at PARC. My main focus was on malware detection using virtualization. The project involved creating a Windows kernel driver that would intercept system calls (like a rootkit) on the guest operating system, and then reporting back the state of the guest to the host. Additional work focused on writing security mechanisms to protect code running under a virtual machine.

Professional Experience

Developer
Tovaris: The Digital Identity Company
2000-2001

I worked part time doing development in C++ for the Mithril Secure Server (an encrypted email solution). I mostly wrote CGI code for administering the servers from a front-end, although I did do some work on the back-end. This involved getting very familiar with the OpenSSL libraries. Most of the development was done under OpenBSD, using C++, though I also did some work in Perl.

Technical Support / Developer / System Administrator
Broadband Network Services, Inc.
1999-2000

I handled all of the technical support questions via telephone and e-mail. I maintained and administrated all of our databases using MySQL. This included setting up new database customers, adding and removing databases, and maintaining MySQL. I used PHP, Perl, and bash to write scripts to aid in system administration and to automate other common tasks. I handled most of the website development that we were hired to do; this included writing scripts, HTML, and database management. My administrative responsibilities included maintaining our primary and secondary DNS, Sendmail, Apache, and PHP. I also aided in creating and removing accounts, setting up new virtual hosts, setting up and maintaining network monitoring, and maintaining hardware; this included building and configuring computers.

Teaching Experience

Information Security & Privacy (46-861)
Carnegie Mellon University
Fall 2007

Teaching assistant duties included developing course materials (topics for lectures, assignments, and exams), grading assignments and exams, holding office hours, and mentoring students about semester-long projects.

Computers and Society (15-290)
Carnegie Mellon University
Spring 2006

Teaching assistant duties included giving guest lectures, creating assignments and exams, grading assignments and exams, holding office hours, and mentoring students about semester-long projects.

Information Security (CS 451)
University of Virginia
Fall 2003

Teaching assistant duties included giving guest lectures, creating assignments and exams, grading assignments and exams, and holding office hours.

Intellectual Property (TCC 200)
University of Virginia
Fall 2003

Teaching assistant duties included grading assignments and holding office hours.

Advanced Software Development Methods (CS 340)
University of Virginia
Spring 2003, Spring 2004

Teaching assistant duties included grading assignments and exams, and holding office hours.

Engineering Software (CS 201J)
University of Virginia
Fall 2002

Teaching assistant duties included grading assignments and holding office hours.

Research Grants

- Google Faculty Research Award, *Designing Usable Certificate Dialogs in Chrome*. Principal Investigator, 2010. Budget: $60,000.

- NSF Trustworthy Computing, Small, *Interfaces to Reduce Human Error in Access Control Policy Authoring*. Principal Investigator (Co-PIs: Shriram Krishnamurthi and Kathi Fisler), 2010. Budget: $500,000; *Recommended for funding, though upon accepting a job within the government, we were forced to subsequently withdraw the proposal.*

Professional Activities

- **Program Committees**
  2013: CHI; Symposium On Usable Privacy and Security (SOUPS)
  2012: Symposium On Usable Privacy and Security (SOUPS); New Security Paradigms Workshop (NSPW)
  2011: Symposium On Usable Privacy and Security (SOUPS); New Security Paradigms Workshop (NSPW); Computers, Freedom, and Privacy (CFP) Conference (poster session co-chair); Software and Usable Security Aligned for Good Engineering (SAUSAGE) Workshop (co-chair)
  2010: Symposium On Usable Privacy and Security (SOUPS)
  2008: Conference on Information and Knowledge Management (CIKM)
  2007: CHI 2007 Workshop - Security User Studies: Methodologies and Best Practices; Anti-Phishing Working Group eCrime Researchers Summit (poster session co-chair)
  2006: Computers, Freedom, and Privacy (CFP) Conference
- **Standards Committees**
  2007-2008: W3C Web Security Context (WSC) Working Group
  2004-2006: W3C Platform for Privacy Preferences (P3P) 1.1 Working Group
- **Leadership Roles**
  Legislative Concerns Chair, Board of Directors
  National Association of Graduate and Professional Students (NAGPS), 2006-2008
  Vice President for External Affairs
  Carnegie Mellon Graduate Student Assembly, 2006-2008

Awards and Nominations

- **ISR Best Published Paper**, 2012

  The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study, received the Best Published Paper Award at the 2012 INFORMS Conference (with J. Tsai, L. Cranor, and A. Acquisti).
- **SOUPS Best Paper Award**, 2012

  Android Permissions: User Attention, Comprehension, and Behavior, received the Best Paper Award at the Symposium on Usable Privacy and Security (with A. P. Felt, E. Ha, A. Haney, E. Chin, and D. Wagner).
- **CHI Best Paper Nominee**, 2011

  Of Passwords and People: Measuring the Effect of Password-Composition Policies, received an honorable mention at CHI 2011 (with with S. Komanduri, R. Shay, P. G. Kelley, M. Mazurek, L. Bauer, N. Christin, and L. F. Cranor).
- **CHI Best Paper Nominee**, 2008

  You've Been Warned: An Empirical Study on the Effectiveness of Web Browser Phishing Warnings, received an honorable mention at CHI 2008 (with L. Cranor and J. Hong).
- **Tor Graphical User Interface Design Competition**, 2006

  Phase 1 Overall Winner (with L. Cranor, J. Hong, P. Kumaraguru, C. Kuo, S. Romanosky, J. Tsai, and K. Vaniea).
- **University of Virginia Dean's List of Scholars**

  I was included on the Spring 2003 and 2004 Dean's List of Scholars.
- **Publisher's Clearing House Finalist**

  I *may* already be a winner.

Last modified February 2013.

EXHIBIT A-2

Expert Witness Report

In re: LinkedIn User Privacy Litigation,
No. 12-cv-03088-EJD (N.D. Cal.)

Serge Egelman, Ph.D.

April 30, 2013

Contents

1. Introduction

In this report, I show that decades of research have established that consumers factor privacy considerations into their purchasing decisions. Among these privacy considerations is the secure storage of their personal information. Awareness of these considerations has driven businesses to both improve and disclose their information security practices in an effort to attract consumers. LinkedIn's paying customers believed that having their account information securely stored was just one of the many benefits of paying a subscription fee. They expected that part of their subscription fee was going towards the secure storage of their personal information using practices that met or exceeded industry standards.

Due to weaknesses discovered in the SHA-1 hashing algorithm [47], industry experts and government agencies have recommended against using SHA-1 [39, 38, 32], since at least 2005. Likewise, adding a "salt" to hashed passwords has been an industry standard since the late 1970s [31]. Thus, LinkedIn's use of unsalted SHA-1 hashes did not follow industry standards. After LinkedIn's data breach, they received significant media attention regarding their substandard security practices. This breach was far more extensive than LinkedIn has admitted. Based on previous incidents and existing research, had LinkedIn disclosed their actual security practices, their customers would have been made aware. Once made aware, existing paying customers likely would have demanded refunds and terminated their subscriptions. However, because LinkedIn did not disclose their actual practices, their customers were not made aware, and therefore did not receive what was expected.

These facts indicate that LinkedIn's paying customers did not receive all of the benefits that they believed they were purchasing through their monthly subscription fees.

2. Consumers Demand Information Security

In this section I show that it is well established in the academic literature that consumers demand privacy and take measures to achieve it, such as by paying premiums for increased information security. Merchants are aware of these demands and offer premium services in response to them.

I present the results of a survey that I designed and deployed that show that LinkedIn's paying customers expected a greater degree of information security than their non-paying customers; paying customers expect that a certain portion of their subscription fee is financing the secure storage of their personal information. At a bare minimum, these customers expect that their information will be protected using industry standards. Likewise, I show that these expectations correlate with

the amount of the subscription price, both in terms of the expected information security practices, as well as the resulting anger when it is discovered that substandard practices are being used. Finally, I present the results of a second survey that I designed and deployed that show that when given the choice of joining a website with poor information security practices, most users will refuse to sign up because they view these poor practices as negatively impacting the utility of the entire website. This suggests that had LinkedIn disclosed their actual practices, they would have received many fewer paying customers.

2.1. Research on Privacy Preferences & Behaviors

Consumer privacy expectations have been studied for several decades and one clear conclusion is widely known in the academic community: people care about privacy and take measures to ensure that their information is not handled inappropriately. One very important aspect of privacy, and the subject of this case, is the secure storage of personal information.

Westin performed a series of consumer surveys between 1978 and 2004 to examine the proportion of consumers who were concerned with privacy [25].[1] Across all of his studies, he classified the public into three groups: privacy fundamentalists, privacy pragmatists, and the privacy unconcerned. Privacy fundamentalists distrust organizations with whom they share personal information and go to great lengths to prevent their personal information from being disseminated, sacrificing specific benefits in order to better control their privacy. Privacy pragmatists make calculated decisions that balance control of their privacy with other benefits. The privacy unconcerned are generally willing to trade privacy for other benefits and trust that organizations will handle their data in a responsible manner. Westin observed that those who do not consider privacy are in the minority: the privacy unconcerned made up 18% of the public in 1991 [48], 16% of the public in 1996 [50], 13% of the public in 1998 [49], 20% of the public in 2001 [22], and 10% of the public in 2003 [40]. Surveys conducted by other researchers have corroborated these results (e.g., [1, 6]). Thus, it is well known that over 80% of the public values privacy and considers it when making purchasing decisions.

Other research has shown that once consumers are made aware of the secondary uses of their information, they raise "strenuous objections" [46]. In response to the public demand for privacy, in 1998, the U.S. Federal Trade Commission (FTC) outlined five principles for online merchants to follow [45]:

[1] Less research has been performed since 2004 to establish whether people value privacy, since it is widely understood that they do. Research has since shifted to examine the extent to which they value it, when balanced with other concerns, and how this changes based on specific circumstances.

1. **Notice**: Organizations should provide their customers with the details of their privacy practices.
2. **Choice**: Consumers should have the ability to decline certain uses of their information.
3. **Access**: Consumers should have the ability to view the information that an organization has collected about them.
4. **Security**: Organizations should store consumer data in a secure manner to prevent its unauthorized dissemination or corruption.
5. **Redress**: If an organization deviates from its stated policy, consumers should have access to enforcement mechanisms.

I emphasize that the secure storage of personal data is an integral aspect of privacy. Not only is this recognized in the U.S. through the FTC's privacy guidelines, but it is also recognized internationally. Since 1980, the Organization for Economic Co-operation and Development (OECD) has issued privacy guidelines. Among their set of eight principles is the "Security Safeguards Principle," which describes the secure storage of personal data [33]:

> Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification, or disclosure of data.

Based on the plethora of available research and privacy guidelines, it is widely known among businesses that consumers are willing to pay increased prices in order to do business with merchants who better protect their privacy by following these guidelines. For instance, in 2002, Rust et al. noted the market for privacy-related services that had arisen due to consumer demand [37]: "They essentially provide cloaking services to consumers through untraceable pseudonymous identities at a monthly service charge of $50 to $60. Given that ISPs generally charge a $20 monthly service fee, consumers essentially pay $30 to $40 per month for the privacy-anonymity feature."

Researchers have taken these observations a step further by performing controlled experiments to examine the extent to which consumers are willing to spend additional money in exchange for increased privacy protections (e.g., [2, 3, 19, 8, 7]). In my own research over the past decade, I have performed several laboratory experiments to examine consumers' willingness to pay for privacy as part of larger online purchasing decisions (i.e., the extent to which they were willing to pay privacy premiums) [16, 11, 43, 10]. In all of my experiments, I observed that a sizable percentage of consumers wish to pay higher prices in exchange for stronger privacy

protections, which includes the secure storage of their personal information. The corollary to this is that when consumers pay for premium services, they expect to receive increased privacy protections.

2.2. Expectations of Social Networking Users

During the week of April 1, 2013, I designed and deployed a survey to examine how paying users of social networking websites expect their personal data to be handled relative to non-paying users. I recruited 506 respondents using Amazon's Mechanical Turk, 65% of whom were male, 34% were female, and 1% declined to state their gender. Survey respondents were located in the U.S., and ranged in age from 18 to 66 ($M = 26.5$). So as not to bias respondents, I did not ask them about LinkedIn's data breach, but instead asked them to imagine they had created profiles on a fictitious social networking website, "Site A." I primed them to think about the following data that would be contained in their profiles: name, education history, employment history, interests, and connections (e.g., friends, coworkers, family members, etc.). This is the same data that might appear in a user's LinkedIn profile.

In the first section of the survey, I asked respondents to imagine that this fictitious social networking website offered both free and paid accounts. I asked them to indicate how they would expect a paid account to differ from a free account. To not bias them towards one particular response, I provided a free-response text box, as well as the following multiple choice options presented in random order:

- Fewer advertisements
- Better security (protection of account information)
- Better customer service
- More features (e.g., tools to help you connect with others)

Overall, I observed that 215 respondents (42.5% of 506) expected that a paid account would feature better security. This was the second most popular answer after "fewer advertisements." Since I was not directly asking respondents about security, this proportion likely represents a lower bound.

In the second section of the survey, I directly asked respondents to compare a free account with a paid account, with regard to the protection of personal information. In order to be able to control for the amount of the subscription price, I randomly assigned respondents to five different between-subjects conditions. Each condition differed based on the instructions displayed at the top of the page:

1. Imagine that you have a free account on Site A. This means that the company relies on advertisements targeted based on your interests.

2. Imagine that you pay $25/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

3. Imagine that you pay $50/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

4. Imagine that you pay $75/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

5. Imagine that you pay $100/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

I observed that respondents who were assigned to one of the four paid account conditions believed that their personal information would be stored significantly better than the account information of non-paying users. This finding was a result of performing a one-sample Wilcoxon Signed Rank test ($p < 0.0005$) to compare the expected median value of 3.0 (i.e., that security practices would be "the same") with the collected data. Respondents reported a median value of 4.0, which means that they believed that their personal information would be stored "better" than non-paying users' personal information. These results stood in contrast with the respondents assigned to the non-paying account condition, whose observed median response did not significantly differ from the expected median value of 3.0 ($p < 0.441$). What this means is that users who pay a subscription fee for an online account expect that their personal information will be protected significantly better than non-paying users' personal information, whereas non-paying users do not believe that they are missing out on anything.

I also asked respondents to report how well they believed the fictitious website would store their personal information relative to other free social networking websites. Responses were reported using a 5-point Likert scale ranging from "much worse" (1) to "much better" (5), with "the same" (3) as the neutral option. Corroborating the previous finding, respondents assigned to the paid conditions reported significantly higher values than respondents assigned to the non-paying condition ($U = 8345.0$, $p < 0.0005$; Mann-Whitney U test). Respondents in the paying conditions reported a median of 4.0 (i.e., "better" than the other free social networking website), whereas respondents in the non-paying condition reported a median of 3.0 (i.e., "the same" as the other free social networking website).

2.3. Privacy Expectations Correlate with Price

In the survey that I deployed, I also noted that respondents' privacy expectations were correlated with the prices they believed they would be paying. I performed a Pearson correlation to examine the relationship between the randomly assigned price conditions and the care with which respondents believed their data would be protected relative to other free websites. This correlation was both positive and statistically significant ($r = 0.297$, $p < 0.0005$). What this means is that the more users pay for a website subscription, the better they believe their information will be protected relative to users of a free website.

Finally, I asked respondents to imagine that their passwords on this fictitious website had been compromised because the site had been using substandard security practices. I asked them to report how angry they would be using a 7-point Likert scale, from "unconcerned" (1) to "angriest" (7). I performed another Pearson correlation and observed that respondents' anger levels were significantly correlated with the assigned prices ($r = 0.157$, $p < 0.0005$). That is, the more they believed they were paying for a subscription to this website, the angrier they were at the use of substandard security practices.

2.4. Social Networking Users Demand Security

During the week of April 22, 2013, I designed and deployed an additional survey to examine how social networking website users evaluate data security as part of the overall utility of a social networking website. I recruited 630 respondents using Amazon's Mechanical Turk,[2] 65% of whom were male, and 35% were female. Survey respondents were in the U.S. and ranged in age from 18 to 70 ($M = 26$). In the first part of the survey, I presented survey respondents with a description of a fictitious social networking website, "Site A." I included a description of the website's features:

- Search tools to meet new people
- Communication tools to stay in touch with friends/coworkers
- Profile information to share your interests and employment history
- Ability to see who's viewed your profile

Included among these features was one of two randomly-assigned descriptions of the website's security practices:

[2]Denominators in this analysis do not always add up to 630 because some respondents did not answer every question.

1. Your personal data is protected using industry standard security practices

2. Limited protection of your personal data (i.e., below industry standards)

I additionally included another between-subjects condition for the monthly price of this fictitious website:[3]

1. **Imagine that an account on Site A is free.** This means that the company relies on advertisements, targeted based on your interests.

2. **Imagine that Site A charges $25/month for an account.** This means that the company also stores your billing information (i.e., address and credit card number).

3. **Imagine that Site A charges $50/month for an account.** This means that the company also stores your billing information (i.e., address and credit card number).

4. **Imagine that Site A charges $75/month for an account.** This means that the company also stores your billing information (i.e., address and credit card number).

5. **Imagine that Site A charges $100/month for an account.** This means that the company also stores your billing information (i.e., address and credit card number).

I asked respondents to report how useful they believed this website to be using a 5-point Likert scale, from "extremely useful (5)" to "not at all useful (1)." I performed a Mann-Whitney U test to compare the two security conditions and observed a statistically significant difference ($U = 43626.5$, $p < 0.026$): respondents who believed that the site used industry standard security practices viewed it as having much higher utility. Respondents who were told that the site used substandard security practices reported a median utility of 1.0, or "not at all useful."

I included an open-ended question to determine, unprompted, why or why not participants would choose to not join this fictitious website. I examined the 480 responses from respondents who were told that the site charged for membership and indicated that they would not create accounts, to examine the primary reasons for their refusals. When respondents were told that the website used substandard security practices, I observed that significantly more reported privacy concerns as

[3]Thus, subjects were randomly assigned to one of ten between-subjects conditions: two possible security conditions and five possible price conditions.

their primary reason for choosing not to subscribe ($p < 0.0005$, Fisher's exact test): 32% reported privacy concerns in the sub-standard security condition, whereas only 4% reported privacy concerns in the industry-standard security condition. Some example quotes include:

- *"If I'm paying $25 a month I would hope to expect a larger amount of protection above industry standards."*
- *"I would expect a company that charges any amount of money to protect my data to the best of industry standards."*
- *"I will not pay $100/month to a site that doesn't even promise to keep my personal information as safe as other companies do."*
- *"...not only am I risking my information being exposed, but I am paying for it."*
- *"Even if the website was free, the 'Limited protection of your personal information (i.e., below industry standards)' would make me not want to sign up."*
- *"I care deeply about the protection of my personal data."*
- *"...if there is limited protection of personal data, that is the most important reason not to join."*
- *"I would not pay $50/month for limited protection of my personal data."*
- *"I would not pay to join a social networking site, especially one with terrible data protection."*
- *"I would not pay for a site that has limited protection of my personal data..."*
- *"It costs money and it gives me limited protection, sounds terrible."*
- *"...50/month is absurdly high of a price for something that offers limited protection of my personal data."*

In the second part of the survey, I directly asked respondents to choose between two different social networking websites, "Site B" and "Site C." I used the same description as I used for the previous fictitious social networking website, however, I randomly labeled one as using industry standard security practices, whereas the other was randomly labeled as offering limited security protections (i.e., below industry standards). I included the same five randomly-assigned price conditions that were used in the first half of the experiment, such that both websites were

labeled as costing the same amount (i.e., free, $25/month, $50/month, $75/month, or $100/month). I asked participants to select the website that is likely to be most useful to them, as well as the website to which they would ultimately subscribe, if they were forced to pick one. All values were reported using a 5-point Likert scale: "Site B (5)," "Likely Site B (4)," "Either Website (3)," "Likely Site C (2)," or "Site C (1)."

When Site B was labeled as following industry standard security practices, 252 (82% of 308) respondents indicated that it was the most useful website or was likely the most useful; only 9 respondents selected "Site C" or "Likely Site C" (3% of 308). Whereas when Site C was labeled as following industry standard security practices, 235 (74% of 319) respondents indicated that Site C was the most useful or was likely the most useful. A Mann-Whitney U test indicate that this contrast is statistically significant ($U = 4981.5$, $p < 0.0005$). Thus, when calculating a website's utility, the vast majority of respondents heavily factored in each website's stated security practices.

Likewise, this finding was corroborated when choosing a website subscription; respondents were significantly more likely to choose Site B when it was labeled as following industry standard security practices ($U = 2785$, $p < 0.0005$). Across both conditions, only 20 (3% of 629) respondents selected the website that was labeled as following substandard security practices.

Thus, these survey results show that users heavily weigh a website's security practices as part of its perceived utility. Users actively avoid websites that are perceived to follow security practices that are below industry standards.

3. Standard Practices for Password Storage

In this section I provide background information on industry standards for secure password storage. I will explain the process of "hashing" a password to transform it into an unreadable format, and why hashing alone provides insufficient security; it has been standard practice since the 1970s to both hash and "salt" passwords [31]. LinkedIn stored users' passwords using unsalted SHA-1 hashes, which violates industry standards for two reasons: the use of unsalted hashes and the use of the outdated SHA-1 algorithm.

3.1. Hash Functions

Industry standards dictate that account passwords be stored in a "hashed" format, which means that a one-way function is applied to them. This is done for two reasons: to prevent malicious insiders from viewing users' passwords (e.g., a system administrator who has legitimate access to the password database) and to limit

the damage in the event that the database is breached (i.e., the revealed hashes do not reveal the underlying plaintext passwords). When a user attempts to log in to a system, the password she types is transformed with the same hash function as the stored passwords. Next, the hashed value is compared with the hashed value in the database. If the two match, that means that she typed the correct password and the system authenticates her.

Hash functions are designed to reveal no information about the underlying input (i.e., the plaintext password). Specifically, minor changes to the input data should result in drastic changes to the resulting hashes. For instance, consider the SHA-1 hashes of the words "password" and "passwords:"

$$\text{SHA1}(\text{“password”}) = \texttt{5baa61e4c9b93f3f0682250b6cf8331b7ee68fd8}$$

$$\text{SHA1}(\text{“passwords”}) = \texttt{a267f7dba707256b0b664dee86ab9ae8b4941218}$$

Despite a Levenshtein distance of 1 between the two input strings to the hash function,[4] the resulting outputs have a Levenshtein distance of 36.

Because a hash is a one-way function, the way in which an attacker guesses a hashed password is by repeatedly applying the same hashing algorithm to guesses until a resulting hash matches the hash of the targeted password. This means that a hash can fail in one of two ways: an attacker can use cutting edge technology to compute billions of hashes in a reasonable amount of time, or the attacker can consult a large corpus of pre-computed hashes.

To prevent an attacker from computing large numbers of hashes, it is desirable for hashing algorithms to be computationally expensive to compute (i.e., each operation should take a reasonable amount of time). For instance, if it takes the computer 100ms to compute a hash, this would hardly be noticeable to a legitimate user when authenticating. However, an attacker who attempts to guess every possible 8-character combination will need to make 95^8 possible guesses.[5] It would therefore take this attacker over 21 million years. Thus, the choice of hashing algorithm has a profound impact on the overall security of a system, and industry standards slowly evolve to recommend different hashing algorithms based on the speed of current technology (i.e., as computers get faster, industry standards recommend using slower and more complex hashing algorithms).

With regard to SHA-1, the hashing algorithm used by LinkedIn, the 100ms per hashing operation used in the previous example is a vast overstatement: specialized technology has been shown to perform up to 63 billion SHA-1 hashes

[4] The Levenshtein distance is the minimum number of insertions, deletions, or substitutions required to transform one text string into another [26].

[5] A standard keyboard contains 95 possible characters: 26 lowercase letters, 26 uppercase letters, 10 digits, and 33 symbols.

per second [35]. This means that SHA-1 hashes for every possible 8-character password can be generated in under thirty hours. But even without specialized hardware, publicly available cloud computing platforms (e.g., Amazon's EC2 or Microsoft's Azure) and commodity hardware allow anyone to compute billions of SHA-1 hashes per second. This is part of the reason why the National Institute of Standards and Technology (NIST) recommended that government agencies stop using SHA-1 in 2006 [32]:

> "Federal agencies should stop using SHA-1 for digital signatures, digital time stamping and other applications that require collision resistance as soon as practical, and must use the SHA-2 family of hash functions for these applications after 2010."

Another way of cracking passwords is through the use of "rainbow tables." A rainbow table is a list of input strings and their resulting hashes that have been precomputed, in order to save someone the time of computing the hashes themselves; a rainbow table is simply a lookup table that an attacker can use to avoid computing their own hashes. Thus, cracking a password using a rainbow table involves looking up the targeted hash and seeing what the matching input string is (i.e., the password). Rainbow tables can be freely downloaded online. Some websites even allow users to query hashes in a fraction of a second, without the need to download the tables themselves. It took me less than a minute to find a web-based rainbow table lookup tool that claims to query over 15 billion SHA-1 hashes in realtime. Thus, simply hashing passwords is no longer effective, and industry standards recommend "salting" passwords to prevent attacks using rainbow tables.

3.2. Salting

The use of rainbow tables can be mitigated by increasing the number of possible hashes, such that the resulting rainbow tables would be so big that they would be infeasible to store. This is done by "salting" the passwords prior to hashing, a process that has been recommended since the 1970s [31].

Salting is the process of concatenating a plaintext password with a series of randomly generated characters prior to hashing. The salt is then stored alongside the hash. Salting is performed for two reasons. First, salting prevents two identical passwords from yielding the same hash. It is well known that users frequently choose passwords in predictable ways, and therefore many users are likely to have the same password [13]. If hashes are properly salted, an attacker will not know how many hashes correspond to the same password, and will therefore have to expend the same amount of effort to compromise each individual account.

Whereas if passwords are not salted and the attacker is interested in compromising as many accounts as possible (rather than targeting a specific user), she will go after the hash that occurs most frequently because once cracked, she now knows the password for every account with this same hash.

In addition to preventing identical passwords from yielding the same hashes, salting is also performed to render rainbow tables ineffective. This is because the rainbow tables would need to include both the password and the salt, which therefore substantially increases the total number of hashes. For example, if each 8-character password is concatenated with a 2-character random salt, this would effectively create a 10-character password. The addition of the salt would increase the size of the rainbow table by a factor of 9,025 (i.e., 95^2). Thus, this rainbow table would need to be over 5,000TB, which is simply impractical for an independent attacker (i.e., only governments and very large corporations can afford this amount of storage space). With an even longer salt, as is generally recommended, rainbow tables are rendered entirely useless because they cannot be easily stored or transferred. Thus, it has been longstanding practice to both salt and hash stored passwords.

4. LinkedIn's Password Storage Practices

In this section, I show that far from following industry standards, LinkedIn's practices were so outdated that they received a significant amount of media attention. LinkedIn used a deprecated algorithm, SHA-1, to hash their customers' passwords, and also chose not to salt them. This media attention did not focus on the fact that they had been hacked, which is a relatively common occurrence, but that they had not been following commonly accepted industry standards that would have mitigated the harm. Furthermore, the password breach was likely far worse than LinkedIn has heretofore admitted.

Finally, I show that had LinkedIn disclosed their actual practices in their privacy policy, rather than making erroneous statements about following industry standards, the public would have been made aware of their substandard practices.

4.1. Response to LinkedIn's Practices

Websites are under constant attack by hackers. While websites take many precautions to ward off potential intruders, it is unlikely that they will be able to detect and patch all vulnerabilities before one gets exploited. Therefore, that a website gets hacked is no longer newsworthy by itself. However, many precautions can still be taken to both detect potential intruders and limit the amount of damage that they may cause. What is different about the LinkedIn case, compared to the tens

of thousands of other websites that are hacked every year, is that LinkedIn failed to take acceptable measures to mitigate the damage.

The popular media was quick to condemn LinkedIn's data breach not because it occurred, but because LinkedIn did not take the expected steps to prevent it. As one article put it, "the site was flat-out doing security wrong" [17]. Other relevant quotes include:

- "What has surprised customers and security experts alike is that a company that collects and profits from vast amounts of data had taken a bare-bones approach to protecting it. The breach highlights a disturbing truth about LinkedIn's computer security: there isn't much" [34].
- "LinkedIn's loss of 6.5 million passwords is bad enough, but the fact they were easily deciphered shows a stunning lack of care for software security" [18].
- "Is it too much to ask for a company like this to take security seriously enough to do a better job protecting and securing their users' passwords?" [24].
- "[W]e have yet to find out why nobody objected to them protecting 150+ million user passwords with 1970s methods" [23].
- "[T]here's no good excuse for a site this prominent to not have a salted, secure password hashing system" [20].
- "The site suffered a break-in, and the intruders swiped files containing many users' logins and passwords. That's not good, but it's a setback that could have been mitigated by following some longstanding best practices, like encrypting that data so that even if someone should steal it, they couldn't make any sense of it. But apparently that practice had not been followed" [17].
- "The only thing worse they could have done would be to put straight passwords in a file, but they came pretty close to that by failing to salt" [30].
- "For security gurus, this is kind of like How to Protect Users 101" [36].

4.2. The Extent of the Breach

Of the 6.5 million hashes that were made public, all 6.5 million were unique (i.e., there were 6.5 million different hashes with no duplicates). One would expect to see this if LinkedIn had salted their hashes, because the random salt would ensure that even users who chose the same password would each have unique hashes. However, because the LinkedIn passwords were not salted, and because the 6.5 million publicized hashes are unique, this indicates that many more accounts

were likely compromised. Specifically, the disclosed 6.5 million unique hashes correspond to 6.5 million unique passwords.

Based on all the research on password composition behaviors that has been performed to date, it is a near impossibility that 6.5 million users chose 6.5 million unique passwords. For instance, Bonneau observed that of 70 million Yahoo! passwords, roughly half were unique (i.e., half the users had chosen the same password as other users) [4, 5]. Likewise, Malone and Maher found that of over 32 million RockYou passwords, roughly 44% were unique [28]. Thus, LinkedIn's publicly available list of 6.5 million unsalted hashes corresponds to many more compromised accounts, likely a factor of two (estimating from prior password breaches). However, based on the existing research on why user data is stolen and how online criminals operate, I believe that this is still a lower bound for the number of accounts that were compromised.

Websites are hacked so that users' personal information can be traded as a commodity. A vast underground economy exists wherein these illicit goods are bought and sold [41, 14, 15]. When information is sold on the black market, buyers are cautious due to the unregulated nature of this market. As such, sellers generally provide free samples of the "goods" in order to demonstrate that they are legitimate. Since information is sold in bulk, this generally means disclosing a small subset of stolen account credentials or personal details [14]; an attacker has little incentive to publicize stolen data unless it is to facilitate a much larger future sale. Thus, the publication of 6.5 million passwords is likely only a small subset of what was stolen: the attacker very likely also has a much larger set of passwords, as well as their associated usernames. From my professional experience, I believe that the attacker disclosed 6.5 million hashes to demonstrate the legitimacy of a much larger set that he or she was attempting to sell. The evidence suggests that the attacker released a subset of unique hashes, rather than the entire stolen set or even a random sample of the entire set (i.e., even a random sample of 6.5 million unsalted hashes would yield duplicates).

5. Dissemination of Website Privacy Practices

In this section, I summarize prior research to show that users learn about poor privacy and security practices via word of mouth and popular media coverage. From this research, it is obvious that had LinkedIn stated their actual password storage practices rather than misrepresenting them, their users would have discovered this.

The contents of privacy policies are frequently disseminated by a small group of experts after they discover that a particular policy fails to meet expectations or greatly diverges from industry standards. For instance, the popular photo-sharing website, Instagram, has a privacy policy that is over 3,000 words and requires over

twelve years of formal education to be understood [21],[6] which exceeds the average privacy policy's complexity [29]. Despite its complexity, when Instagram changed the policy to include a clause saying that users' photos will be used for advertising purposes, the popular media quickly noted the change (e.g., [42]). Users suddenly became aware of the change and cried foul because it did not comply with their expectations. Worried that they would lose users, Instagram subsequently removed the offending clause from their privacy policy and launched a public relations campaign to inform their users of the change [44].

In my own recent research on how Facebook users share personal information with third-parties, I have observed that despite the complexity of privacy disclosures, many users become aware of privacy practices through word of mouth, popular media accounts, and prior interactions [9]. In another study, I observed that users become aware of inappropriate data usage by smartphone applications through application reviews and word of mouth, rather than through primary means (i.e., application developers' disclosures) [12]; users become upset when they invariably discover, through secondary sources, that their information is being used in ways that defy their expectations [27].

Based on the research on how users become aware of privacy policies, had LinkedIn's privacy policy disclosed that they were using substandard practices or specifically mentioned the use of unsalted SHA-1 hashes, it is clear that this information would have been quickly disseminated to the public. In my professional opinion, LinkedIn's existing users would have learned about their substandard practices through word of mouth and popular media accounts and likely would have been outraged; existing paying customers likely would have discovered that they were not receiving the protections for which they believed that they were paying, through their monthly subscription fees, and potential new customers likely would choose not to subscribe.

6. Conclusion

Decades of privacy research have established that consumers desire privacy and are willing to pay higher prices to achieve it. Merchants are aware of this and publicize their use of strong security practices as a way of demonstrating value to potential customers. Customers are attracted to these premium services because they do not want their personal information to be inappropriately disclosed. LinkedIn's customers are no different: I commissioned a survey of existing social networking website users who indicated that when paying for a website subscription, they expected that their information would be protected using industry standard practices

[6] Calculated using the Flesch-Kincaid Grade Level test.

or better. In a second survey, I showed that users factor information security practices into their perceptions about the utility of a website, and will avoid websites that offer substandard security practices.

In this report, I showed why storing password data as unsalted SHA-1 hashes—the way in which LinkedIn stored their paying customers' passwords—does not conform to industry standards. Unlike other website security breaches, LinkedIn's garnered significant media attention due to their use of these outdated practices. Had LinkedIn chosen to disclose their actual password storage practices in their privacy policy, prior research shows that their customers would have become aware. After being made aware of LinkedIn's true practices, customers would have immediately discovered that they were not receiving the level of information security for which they believed they were paying. As a result, LinkedIn's existing paying customers would have demanded refunds and/or canceled their subscriptions, while new potential customers would have declined to pay the subscription price. However, LinkedIn chose to misrepresent their practices, so their customers were not made aware until their data had already been stolen.

References

[1] Ackerman, M. S., Cranor, L. F., and Reagle, J. Privacy in e-commerce: examining user scenarios and privacy preferences. In *EC '99: Proceedings of the 1st ACM Conference on Electronic Commerce* (New York, NY, USA, 1999), ACM, pp. 1–8. http://www.eecs.umich.edu/~ackerm/pub/99b28/ecommerce.final.pdf.

[2] Acquisti, A., and Grossklags, J. Losses, gains, and hyperbolic discounting: An experimental approach to information security attitudes and behavior. In *Proceedings of The 2nd Annual Workshop on Economics and Information Security (WEIS '03)* (2003).

[3] Acquisti, A., and Grossklags, J. Privacy and rationality in individual decision making. *IEEE Security & Privacy* (January/February 2005), 24–30. http://www.dtc.umn.edu/weis2004/acquisti.pdf.

[4] Bonneau, J. The science of guessing: Analyzing an anonymized corpus of 70 million passwords. In *Proceedings of the 2012 IEEE Symposium on Security and Privacy* (Washington, DC, USA, 2012), SP '12, IEEE Computer Society, pp. 538–552.

[5] Bonneau, J. The science of password guessing. http://www.lightbluetouchpaper.org/2012/05/24/the-science-of-password-guessing/, May 24 2012. Accessed: April 2, 2013.

[6] Cranor, L. F., Reagle, J., and Ackerman, M. S. Beyond concern: Understanding net users' attitudes about online privacy. *AT&T Labs-Research Technical Report TR 99.4.3* (14 April 1999). http://www.research.att.com/resources/trs/TRs/99/99.4/99.4.3/report.htm.

[7] Cvrcek, D., Kumpost, M., Matyas, V., and Danezis, G. A study on the value of location privacy. In *Proceedings of the 2006 Workshop on Privacy in an Electronic Society (WPES'06)* (2006).

[8] Danezis, G., Lewis, S., and Anderson, R. How much is location privacy worth? In *Proceedings of the Workshop on the Economics of Information Security Series (WEIS 2005)* (2005).

[9] Egelman, S. My profile is my password, verify me! the privacy/convenience tradeoff of facebook connect. In *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013).

[10] Egelman, S., Felt, A. P., and Wagner, D. Choice architecture and smartphone privacy: There's a price for that. In *The 2012 Workshop on the Economics of Information Security (WEIS)* (2012).

[11] Egelman, S., Tsai, J., Cranor, L. F., and Acquisti, A. Timing is everything?: the effects of timing and placement of online privacy indicators. In *Proceedings of the 27th international conference on Human factors in computing systems* (New York, NY, USA, 2009), CHI '09, ACM, pp. 319–328.

[12] Felt, A. P., Ha, E., Egelman, S., Haney, A., Chin, E., and Wagner, D. Android permissions: user attention, comprehension, and behavior. In *Proceedings of the Eighth Symposium on Usable Privacy and Security* (New York, NY, USA, 2012), SOUPS '12, ACM, pp. 3:1–3:14.

[13] Florencio, D., and Herley, C. A large-scale study of web password habits. In *WWW '07: Proceedings of the 16th International Conference on the World Wide Web* (New York, NY, USA, 2007), ACM Press, pp. 657–666.

[14] Franklin, J., Paxson, V., Perrig, A., and Savage, S. An inquiry into the nature and causes of the wealth of internet miscreants. In *CCS '07: Proceedings of the 14th ACM conference on Computer and communications security* (New York, NY, USA, 2007), ACM, pp. 375–388.

[15] Geer, D., and Conway, D. What we got for christmas. *IEEE Security & Privacy* (January/February 2008), 88.

[16] Gideon, J., Egelman, S., Cranor, L., and Acquisti, A. Power Strips, Prophylactics, and Privacy, Oh My! In *Proceedings of the 2006 Symposium on Usable Privacy and Security* (July 2006), pp. 133–144.

[17] Hartsock, P. LinkedIn: Unsalted, Assaulted and Faulted. http://www.technewsworld.com/story/75337.html, June 9 2012. Accessed: April 13, 2013.

[18] Hickins, M. LinkedIn Password Breach Illustrates Endemic Security Issue. http://blogs.wsj.com/cio/2012/06/06/linkedin-password-breach-illustrates-endemic-security-issue/, June 6 2012. Accessed: April 13, 2013.

[19] Huberman, B., Adar, E., and Fine, L. Valuating privacy. *IEEE Security & Privacy 3*, 5 (September-October 2005), 22–25.

[20] Information Week. 6.5 Million LinkedIn Password Hashes Leaked. http://www.informationweek.com/aroundtheweb/security/65-million-linkedin-password-hashes-leak/752f72454f7875616a3731383468512b55334f6643673d3d, 2012. Accessed: April 13, 2013.

[21] Instagram. Privacy Policy. http://instagram.com/about/legal/privacy/, January 19 2013. Accessed: April 10, 2013.

[22] Interactive, H. Privacy on & off the Internet: What consumers want. http://www.aicpa.org/download/webtrust/priv_rpt_21mar02.pdf, 2001.

[23] Kamp, P.-H. Linkedin password leak: Salt their hide. *Queue 10*, 6 (June 2012), 20:20–20:22.

[24] Krebs, B. How Companies Can Beef Up Password Security. http://krebsonsecurity.com/2012/06/how-companies-can-beef-up-password-security/, June 11 2012. Accessed: April 13, 2013.

[25] Kumaraguru, P., and Cranor, L. F. Privacy indexes: A survey of westin's studies. Tech. Rep. CMU-ISRI-5-138, Carnegie Mellon University, December 2005. http://reports-archive.adm.cs.cmu.edu/anon/isri2005/CMU-ISRI-05-138.pdf.

[26] Levenshtein, V. I. Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady 10* (Feb. 1966), 707–710.

[27] Lin, J., Sadeh, N., Amini, S., Lindqvist, J., Hong, J. I., and Zhang, J. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (2012), ACM, pp. 501–510.

[28] Malone, D., and Maher, K. Investigating the distribution of password choices. In *Proceedings of the 21st international conference on World Wide Web* (New York, NY, USA, 2012), WWW '12, ACM, pp. 301–310.

[29] McDonald, A., and Cranor, L. The cost of reading privacy policies. In *Proceedings of the Technology Policy Research Conference* (September 26–28 2008).

[30] Mills, E. LinkedIn Confirms Passwords were 'Compromised'. http://news.cnet.com/8301-1009_3-57448465-83/linkedin-confirms-passwords-were-compromised/, June 6 2012. Accessed: April 13, 2013.

[31] Morris, R., and Thompson, K. Password security: A case history. *Communications of the ACM 22* (November 1979), 594–597.

[32] National Institute of Standards and Technology. NIST's March 2006 Policy on Hash Functions. http://csrc.nist.gov/groups/ST/hash/policy_2006.html, March 2006. Accessed: March 28, 2013.

[33] Organization for Economic Co-operation and Development (OECD). OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm, 1980. Accessed: April 28, 2013.

[34] Perlroth, N. Lax Security at LinkedIn Is Laid Bare. *The New York Times* (June 11 2012), B1. http://www.nytimes.com/2012/06/11/technology/linkedin-breach-exposes-light-security-even-at-data-companies.html.

[35] Roberts, P. F. Update: New 25 GPU Monster Devours Passwords in Seconds. http://securityledger.com/new-25-gpu-monster-devours-passwords-in-seconds/, December 4 2012. Accessed: April 2, 2013.

[36] Rowinski, D. Avoiding Password Breaches 101: Salt Your Hash. http://readwrite.com/2012/06/07/avoiding-password-breaches-101-salt-your-hash, June 7 2012. Accessed: April 13, 2013.

[37] Rust, R. T., Kannan, P., and Peng, N. The customer economics of internet privacy. *Journal of the Academy of Marketing Science 30*, 4 (2002), 455–464.

[38] Schneier, B. Cryptanalysis of SHA-1. http://www.schneier.com/blog/archives/2005/02/cryptanalysis_o.html, February 18 2005. Accessed: March 28, 2013.

[39] Schneier, B. SHA-1 Broken. http://www.schneier.com/blog/archives/2005/02/sha1_broken.html, February 15 2005. Accessed: March 28, 2013.

[40] Taylor, H. Most People are "Privacy Pragmatists" Who, While Concerned about Privacy, Will Sometimes Trade It Off for Other Benefits. *The Harris Poll 17* (March 2003). http://www.harrisinteractive.com/vault/Harris-Interactive-Poll-Research-Most-People-Are-Privacy-Pragmatists-Who-While-Conc-2003-03.pdf.

[41] Thomas, R., and Martin, J. The underground economy: Priceless. *USENIX ;login:* (December 2006).

[42] Timberg, C. Instagram, Facebook Stir Online Protests with Privacy Policy Change. http://articles.washingtonpost.com/2012-12-18/business/35908189_1_kevin-systrom-instagram-consumer-privacy, December 18 2012. Accessed: 10, 2013.

[43] Tsai, J., Egelman, S., Cranor, L., and Acquisti, A. The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. *Information Systems Research 22*, 2 (June 2011), 254–268. http://www.guanotronic.com/~serge/papers/isr10.pdf.

[44] Tsukayama, H. Instagram Reminds Users of Privacy Policy Change. http://articles.washingtonpost.com/2013-01-16/business/36384825_1_instagram-function-photos-new-terms, January 16 2013. Accessed: April 10, 2013.

[45] U.S. Federal Trade Commission. Privacy Online: A Report to Congress. http://www.ftc.gov/reports/privacy3/toc.htm, June 1998. Accessed: December 20, 2008.

[46] Wang, P., and Petrison, L. A. Direct marketing activities and personal privacy: A consumer survey. *Journal of Direct Marketing 7*, 1 (1993), 7–19.

[47] Wang, X., Yin, Y. L., and Yu, H. Finding collisions in the full sha-1. In *Proceedings of the 25th annual international conference on Advances in Cryptology* (Berlin, Heidelberg, 2005), CRYPTO'05, Springer-Verlag, pp. 17–36.

[48] Westin, A. F. Harris-equifax consumer privacy survey (1991). Tech. rep., Equifax, Inc., Atlanta, GA, 1991.

[49] Westin, A. F. *E-Commerce & Privacy: What Net Users Want*. Privacy & American Business, Hackensack, NJ, 1998. http://www.pwcglobal.com/gx/eng/svcs/privacy/images/E-Commerce.pdf.

[50] Westin, A. F., and Associates., H. L. . Harris-equifax consumer privacy survey (1996). Tech. rep., Equifax, Inc., 1996.

A. Survey Instrument 1

Social Networking Account Survey

Introduction

In this survey, we will ask you questions about social networking websites and how you use those accounts. Please answer each question as truthfully as possible. Upon completion, you will receive $0.25 via Mechanical Turk.

---

1. Have you created profiles on any of the following social networking websites?

- [ ] Facebook
- [ ] LinkedIn
- [ ] MySpace
- [ ] Twitter
- [ ] Google+

Hypothetical Social Networking Website

Imagine you have created a profile on a new social networking website, Site A. Your profile contains the following information about you:

- Name
- Education history
- Employment history
- Interests
- Connections (e.g., friends, coworkers, family members, etc.)

This new social networking website allows you to keep in touch with acquaintances, meet new people,and share your interests with others.

---

2. This website is debating offering both free and paid accounts. How would you expect a paid account to be different from a free account on this website?

- [ ] Better customer service
- [ ] Better security (protection of account information)

- [ ] Fewer advertisements
- [ ] More features (e.g., tools to help you connect with others)

3. What other differences would you expect between a free and paid account?

Security Practices

For the questions on this page, imagine that this new social networking website, Site A, offers two types of accounts:

- **Free accounts**
- **Paid monthly subscriptions**

Imagine that you have a free account on Site A. This means that the company relies on advertisements, targeted based on your interests.

Imagine that you pay $25/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

Imagine that you pay $50/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

Imagine that you pay $75/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

Imagine that you pay $100/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

4. How well would you expect Site A to protect your personal information relative to other free social networking websites?

Much Better	Better	The Same	Worse	Much Worse
○	○	○	○	○

5. How well would you expect Site A to protect your personal information relative to their customers who pay a monthly subscription fee?

Much Better	Better	The Same	Worse	Much Worse
○	○	○	○	○

6. How well would you expect Site A to protect your personal information relative to free users of Site A?

Much Better	Better	The Same	Worse	Much Worse
○	○	○	○	○

Security Breach

Imagine that you have a free account on Site A. This means that the company relies on advertisements, targeted based on your interests.

Imagine that you pay $25/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

Imagine that you pay $50/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

Imagine that you pay $75/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

Imagine that you pay $100/month to use Site A. This means that the company also stores your billing information (i.e., address and credit card number).

7. Now, imagine that you are informed of a security breach on Site A due to their use of sub-standard security procedures. Hackers capture your password and can access your account. How angry would this make you?

1 (Unconcerned)	2	3	4 (Angry)	5	6	7 (Angriest)
○	○	○	○	○	○	○

Demographics

8. Earlier you mentioned that you have a LinkedIn account. Is your LinkedIn account free or have you upgraded to a paid account?

- ○ I have a free LinkedIn account
- ○ I pay for my LinkedIn account

9. What is your gender?

- ○ Male
- ○ Female

10. What year were you born?

11. What is your occupation?

Thank You!

Thank you for taking our survey. **To receive your payment, please enter the following number on Mechanical Turk: [question("value"), id="28"]**

B. Survey Instrument 2

Social Networking Survey

Introduction

In this survey, we will ask you questions about social networking websites and how you use those accounts. Please answer each question as truthfully as possible. Upon completion, you will receive $0.25 via Mechanical Turk.

---

1. Have you created profiles on any of the following social networking websites?

- [ ] Facebook
- [ ] LinkedIn
- [ ] MySpace
- [ ] Twitter
- [ ] Google+

Membership

For the questions on this page, imagine a new social networking website, Site A. This social networking website offers the following:

- Search tools to meet new people
- Communication tools to stay in touch with friends/coworkers
- Profile information to share your interests and employment history
- Ability to see who's viewed our profile
- Your personal data is protected using industry standard security practices

---

For the questions on this page, imagine a new social networking website, Site A. This social networking website offers the following:

- Search tools to meet new people
- Communication tools to stay in touch with friends/coworkers
- Profile information to share your interests and employment history
- Ability to see who's viewed our profile
- Limited protection of your personal data (i.e., below industry standards)

Imagine that an account on Site A is free. This means that the company relies on advertisements, targeted based on your interests.

Imagine that Site A charges $25/month for an account. This means that the company also stores your billing information (i.e., address and credit card number).

Imagine that Site A charges $50/month for an account. This means that the company also stores your billing information (i.e., address and credit card number).

Imagine that Site A charges $75/month for an account. This means that the company also stores your billing information (i.e., address and credit card number).

Imagine that Site A charges $100/month for an account. This means that the company also stores your billing information (i.e., address and credit card number).

2. How useful would an account on Site A be to you?

Extremely Useful (5)	(4)	Useful (3)	(2)	Not At All Useful (1)
○	○	○	○	○

3. How likely would you be to create an account with Site A?

Extremely Likely	Likely	Unsure	Unlikely	Extremely Unlikely
○	○	○	○	○

4. Please explain why you would or would not create an account on Site A: *

Choose A Website

Imagine that you must choose between joining two different social networking

websites, Site B and Site C.

Site B offers the following features:

- Search tools to meet new people
- Communication tools to stay in touch with friends/coworkers
- Profile information to share your interests and employment history
- Ability to see who's viewed our profile
- Limited protection of your personal data (i.e., below industry standards)

Site C offers the following features:

- Search tools to meet new people
- Communication tools to stay in touch with friends/coworkers
- Profile information to share your interests and employment history
- Ability to see who's viewed our profile
- Your personal data is protected using industry standard security practices

---

Site B offers the following features:

- Search tools to meet new people
- Communication tools to stay in touch with friends/coworkers
- Profile information to share your interests and employment history
- Ability to see who's viewed our profile
- Your personal data is protected using industry standard security practices

Site C offers the following features:

- Search tools to meet new people
- Communication tools to stay in touch with friends/coworkers
- Profile information to share your interests and employment history
- Ability to see who's viewed our profile
- Limited protection of your personal data (i.e., below industry standards)

Imagine that each of these websites offers free accounts. This means that each company relies on advertisements, targeted based on your interests.

Imagine that each company charges $25/month for an account. This means that each company also stores your billing information (i.e., address and credit card number).

Imagine that each company charges $50/month for an account. This means that each company

also stores your billing information (i.e., address and credit card number).

Imagine that each company charges $75/month for an account. This means that each company also stores your billing information (i.e., address and credit card number).

Imagine that each company charges $100/month for an account. This means that each company also stores your billing information (i.e., address and credit card number).

5. Which of the websites listed above is more useful to you?

Site B	Likely Site B	Either Website	Likely Site C	Site C
○	○	○	○	○

6. If you had to choose between joining Site B or Site C, which one would you join?

Site B	Likely Site B	Either Website	Likely Site C	Site C
○	○	○	○	○

Demographics

7. Earlier you mentioned that you have a LinkedIn account. Is your LinkedIn account free or have you upgraded to a paid account?

- ○ I have a free LinkedIn account
- ○ I pay for my LinkedIn account

8. What is your gender?

- ○ Male
- ○ Female

9. What year were you born?

10. What is your occupation?

Thank You!

Thank you for taking our survey. **To receive your payment, please enter the following number on Mechanical Turk: [question("value"), id="28"]**

C. Biography

Serge Egelman is a research scientist in the Computer Science Department at the University of California, Berkeley. His research focuses on online privacy, security, and human-computer interaction. He routinely performs studies to examine the privacy and security practices of Internet users in order to help designers create better systems that have the user in mind. His research topics have included privacy-enhancements to search engines, web browser security warnings, social networking privacy settings, and smartphone application privacy. Dr. Egelman has over thirty peer-reviewed publications, including top conferences and journals. Several systems based on his research have been deployed for hundreds of millions of Internet users.

Serge Egelman received his B.S. in Computer Engineering from the University of Virginia in 2004, and his doctorate from Carnegie Mellon University's School of Computer Science in 2009. Dr. Egelman has performed research at Xerox PARC, Microsoft, Brown University, and the U.S. National Institute of Standards and Technology (NIST).

D. Curriculum Vitae

Serge Egelman

731 Soda Hall
Berkeley, CA 94720
USA

Email: serge@guanotronic.com

Education

- **PhD in Computation, Organizations, and Society**, May 2009
  School of Computer Science, Carnegie Mellon University
- **BS in Computer Engineering**, May 2004
  School of Engineering and Applied Science, University of Virginia

Refereed Journal Publications

- The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. Information Systems Research (ISR), 22(2), June 2011, pp. 254-268 (with J. Tsai, L. Cranor, and A. Acquisti). *Best Published Paper Award!*
- P3P Deployment on Websites. Electronic Commerce Research and Applications (ECRA), Autumn 2008 (with L. Cranor, S. Sheng, A. McDonald, and A. Chowdhury).
- The Real ID Act: Fixing Identity Documents with Duct Tape. I/S: A Journal of Law and Policy for the Information Society, 2(1), Winter 2006, pp. 149-183 (with L. Cranor).

Refereed Conference Papers

- The Importance of Being Earnest [in Security Warnings]. Financial Cryptography and Data Security. 2013 (with S. Schechter), to appear.
- Does My Password Go up to Eleven? The Impact of Password Meters on Password Selection . CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 (with C. Herley, A. Sotirakopoulos, I. Muslukhov, and K. Beznosov), to appear.
- My Profile Is My Password, Verify Me! The Privacy/Convenience Tradeoff of Facebook Connect . CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013, to appear.
- Android Permissions: User Attention, Comprehension, and Behavior. Proceedings of the 2012 Symposium on Usable Privacy and Security (SOUPS). July 2012 (with A. P. Felt, E. Ha, A. Haney, E. Chin, and D. Wagner). *Best Paper Award!*
- Facebook and Privacy: It's Complicated. Proceedings of the 2012 Symposium on Usable Privacy and Security (SOUPS). July 2012 (with M. Johnson and S. Bellovin).
- Oops, I Did It Again: Mitigating Repeated Access Control Errors on Facebook. CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 (with A. Oates and S. Krishnamurthi).
- Of Passwords and People: Measuring the Effect of Password-Composition Policies. CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 (with S. Komanduri, R. Shay, P. G. Kelley, M. Mazurek, L. Bauer, N. Christin, and L. F. Cranor). *Best Paper Nominee!*
- It's All About The Benjamins: An empirical study on incentivizing users to ignore security advice. Financial Cryptography and Data Security. 2011 (with N. Christin, T. Vidas, and J. Grossklags).
- Crying Wolf: An Empirical Study of SSL Warning Effectiveness. The 18th USENIX Security Symposium. 2009 (with J. Sunshine, H. Almuhimedi, N. Atri, and L. Cranor).
- It's No Secret: Measuring the reliability of authentication via 'secret' questions. The 2009 IEEE Symposium on Security and Privacy (with S. Schechter and A.J. Brush).
- It's Not What You Know, But Who You Know: A social approach to last-resort authentication. CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009 (with S. Schechter and R. Reeder).
- Timing Is Everything? The Effects of Timing and Placement of Online Privacy Indicators. CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009 (with J. Tsai, L. Cranor, and A. Acquisti).
- Family Accounts: A new paradigm for user accounts within the home environment. CSCW '08: Proceedings of the 2008 Conference on Computer Supported Cooperative Work. 2008 (with A.J. Brush and K. Inkpen).
- You've Been Warned: An Empirical Study on the Effectiveness of Web Browser Phishing Warnings. CHI '08: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 (with L. Cranor and J. Hong). *Best Paper Nominee!*
- Phinding Phish: An Evaluation of Anti-Phishing Toolbars. NDSS: Proceedings of the ISOC Symposium on Network and Distributed System Security. February 2007 (with Y. Zhang, L. Cranor, and J. Hong).
- An Analysis of P3P-Enabled Web Sites among Top-20 Search Results. Proceedings of the Eighth International Conference on Electronic Commerce. August 2006 (with L. Cranor and A. Chowdhury).
- Power Strips, Prophylactics, and Privacy, Oh My!. Proceedings of the 2006 Symposium On Usable Privacy and Security (SOUPS). July 2006 (with J. Gideon, L. Cranor, and A. Acquisti).

Refereed Workshop Papers

- I've Got 99 Problems, But Vibration Ain't One: A Survey of Smartphone Users' Concerns. The 2nd Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM). October 2012 (with A. P. Felt and D. Wagner).

- How to Ask for Permission. The 7th USENIX Workshop on Hot Topics in Security (HotSec '12). August 2012 (with A. P. Felt, M. Finifter, D. Akhawe, and D. Wagner).
- Choice Architecture and Smartphone Privacy: There's A Price for That. Workshop on the Economics of Information Security (WEIS). June 2012 (with A. P. Felt and D. Wagner).
- How Good Is Good Enough? The Sisyphean Struggle for Optimal Privacy Settings. CSCW 2012 Workshop on Reconciling Privacy with Social Media. February 2012 (with M. Johnson).
- Toward Privacy Standards Based on Empirical Studies. W3C Workshop on Web Tracking and User Privacy. April 2011 (with E. McCallister).
- Please Continue to Hold: An empirical study on user tolerance of security delays. Workshop on the Economics of Information Security (WEIS). June 2010 (with D. Molnar, N. Christin, A. Acquisti, C. Herley, and S. Krishnamurthi).
- Tell Me Lies: A Methodology for Scientifically Rigorous Security User Studies. Workshop on Studying Online Behaviour at CHI'10. April 2010 (with J. Tsai and L. F. Cranor).
- The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. Workshop on the Economics of Information Security (WEIS). June 2007 (with J. Tsai, L. Cranor, and A. Acquisti).
- Security User Studies: Methodologies and Best Practices. CHI '07 Extended Abstracts on Human Factors in Computing Systems. April 2007 (with J. King, R. Miller, N. Ragouzis, and E. Shehan).
- Studying The Impact of Privacy Information on Online Purchase Decisions. Workshop on Privacy and HCI: Methodologies for Studying Privacy Issues at CHI'06. April 2006 (with J. Tsai, L. Cranor, and A. Acquisti).

Book Chapters and Magazine Articles

- Crowdsourcing. To appear in *Ways of Knowing in HCI*, J. Olson and W. Kellogg (Eds.), to be published by Springer (with E. Chi and S. Dow).
- Helping Users Create Better Passwords. *;login:*. December 2012 (with B. Ur, P. G. Kelley, S. Komanduri, J. Lee, M. Maass, M. Mazurek, T. Passaro, R. Shay, T. Vidas, L. Bauer, N. Christin, L. F. Cranor, and J. Lopez).
- Conference Report: SOUPS 2006. *IEEE Security & Privacy*. November/December 2006 (with J. Tsai).
- Conference Report: 14th USENIX Security Symposium. *;login:*. December 2005 (with K. Butler, M. Chow, J. Duerig, B. Hicks, F. Hsu, S. Kelm, and M. Rajagopalan).
- Conference Report: 13th USENIX Security Symposium. *;login:*. December 2004 (with A. AuYoung, E. Cronin, M. Dougherty, R. Greenstadt, S. Kelm, Z. Liang, C. Mano, N. Smith, A. Raniwala, T. Whalen, and W. Xu).
- Suing Spammers for Fun and Profit. *;login:*. April 2004.
- Installation. *Peter Norton's Complete Guide to Linux*. Macmillan Computer Publishing. 1999.
- User Administration. *Peter Norton's Complete Guide to Linux*. Macmillan Computer Publishing. 1999.

Research Experience

Scientist
University of California, Berkeley
September 2011-present

I am currently working with David Wagner's research group to examine privacy and security issues on mobile devices (e.g., smartphones). Specifically, we are examining how users make decisions to install particular applications and how to better alert them to potential malware. We are in the process of creating a new architecture for prompting users when an application requests certain hardware or software abilities.

Scientist
NIST
August 2010-July 2011

I helped design and conduct studies to examine how users interact with authentication systems, specifically password and token-based systems. I co-organized a workshop on the NIST campus to discuss ways in which usable security research and techniques could be formally integrated into the development process, as well as reviewed grant proposals for NIST funding.

Postdoctoral Research Associate
Brown University
August 2009-August 2010

I worked with Shriram Krishnamurthi on creating better interfaces for policy authors to specify access control policies. We conducted studies to determine common policy errors, the causes of these errors, and the types of interfaces that policy authors currently use. We developed a new policy authoring interface that allows users of social networking websites to interactively specify policies in order to more easily detect and clarify ambiguities. We designed and conducted a usability study to validate our tool.

Research Assistant
Carnegie Mellon University
June 2004-May 2009

While pursuing a PhD under the direction of Dr. Lorrie Cranor in the Computation, Organizations, and Society program at CMU, I focused primarily on the usability of privacy and security systems. Areas that I worked in included creating more effective web browser trust indicators, creating usable privacy tools, Internet anonymity, and detection and prevention of phishing attacks. My dissertation is entitled "Trust Me: Designing Trustworthy Trust Indicators." My committee consisted of Lorrie Cranor (chair), Jim Herbsleb, Jason Hong, and Steve Bellovin (Columbia University).

Research Intern
Microsoft Research
July 2008-October 2008

During my second internship at MSR, I conducted two user studies with Stuart Schechter. We first looked at using social networks as a means for authenticating webmail users who had forgotten their passwords. We tested the usability of our

system as well as how susceptible it would be to various attacks. Additionally, I assisted the Internet Explorer team with new designs for their security warnings based on my research. We tested the new warnings in the laboratory using an eye tracker.

Research Intern
Microsoft Research
January 2008-April 2008

I was an intern at MSR working with A.J. Brush and Kori Inkpen on user account models for shared family computers. We examined why the current user account model does not work on computers shared by trusted individuals (i.e. communal home computers) and developed a more appropriate model. I implemented our prototype in C# and ran a usability study. This work was published at the 2008 Computer Supported Cooperative Work (CSCW) conference.

Research Intern
Xerox PARC
June 2006-September 2006

During the summer of 2006, I worked with Jim Thornton in the Computer Science Lab (CSL) at PARC. My main focus was on malware detection using virtualization. The project involved creating a Windows kernel driver that would intercept system calls (like a rootkit) on the guest operating system, and then reporting back the state of the guest to the host. Additional work focused on writing security mechanisms to protect code running under a virtual machine.

Professional Experience

Developer
Tovaris: The Digital Identity Company
2000-2001

I worked part time doing development in C++ for the Mithril Secure Server (an encrypted email solution). I mostly wrote CGI code for administering the servers from a front-end, although I did do some work on the back-end. This involved getting very familiar with the OpenSSL libraries. Most of the development was done under OpenBSD, using C++, though I also did some work in Perl.

Technical Support / Developer / System Administrator
Broadband Network Services, Inc.
1999-2000

I handled all of the technical support questions via telephone and e-mail. I maintained and administrated all of our databases using MySQL. This included setting up new database customers, adding and removing databases, and maintaining MySQL. I used PHP, Perl, and bash to write scripts to aid in system administration and to automate other common tasks. I handled most of the website development that we were hired to do; this included writing scripts, HTML, and database management. My administrative responsibilities included maintaining our primary and secondary DNS, Sendmail, Apache, and PHP. I also aided in creating and removing accounts, setting up new virtual hosts, setting up and maintaining network monitoring, and maintaining hardware; this included building and configuring computers.

Teaching Experience

Information Security & Privacy (46-861)
Carnegie Mellon University
Fall 2007

Teaching assistant duties included developing course materials (topics for lectures, assignments, and exams), grading assignments and exams, holding office hours, and mentoring students about semester-long projects.

Computers and Society (15-290)
Carnegie Mellon University
Spring 2006

Teaching assistant duties included giving guest lectures, creating assignments and exams, grading assignments and exams, holding office hours, and mentoring students about semester-long projects.

Information Security (CS 451)
University of Virginia
Fall 2003

Teaching assistant duties included giving guest lectures, creating assignments and exams, grading assignments and exams, and holding office hours.

Intellectual Property (TCC 200)
University of Virginia
Fall 2003

Teaching assistant duties included grading assignments and holding office hours.

Advanced Software Development Methods (CS 340)
University of Virginia
Spring 2003, Spring 2004

Teaching assistant duties included grading assignments and exams, and holding office hours.

Engineering Software (CS 201J)
University of Virginia
Fall 2002

Teaching assistant duties included grading assignments and holding office hours.

Research Grants

- Google Faculty Research Award, *Designing Usable Certificate Dialogs in Chrome*. Principal Investigator, 2010. Budget: $60,000.

- NSF Trustworthy Computing, Small, *Interfaces to Reduce Human Error in Access Control Policy Authoring*. Principal Investigator (Co-PIs: Shriram Krishnamurthi and Kathi Fisler), 2010. Budget: $500,000; *Recommended for funding, though upon accepting a job within the government, we were forced to subsequently withdraw the proposal.*

Professional Activities

- **Program Committees**
  2013: CHI; Symposium On Usable Privacy and Security (SOUPS)
  2012: Symposium On Usable Privacy and Security (SOUPS); New Security Paradigms Workshop (NSPW)
  2011: Symposium On Usable Privacy and Security (SOUPS); New Security Paradigms Workshop (NSPW); Computers, Freedom, and Privacy (CFP) Conference (poster session co-chair); Software and Usable Security Aligned for Good Engineering (SAUSAGE) Workshop (co-chair)
  2010: Symposium On Usable Privacy and Security (SOUPS)
  2008: Conference on Information and Knowledge Management (CIKM)
  2007: CHI 2007 Workshop - Security User Studies: Methodologies and Best Practices; Anti-Phishing Working Group eCrime Researchers Summit (poster session co-chair)
  2006: Computers, Freedom, and Privacy (CFP) Conference
- **Standards Committees**
  2007-2008: W3C Web Security Context (WSC) Working Group
  2004-2006: W3C Platform for Privacy Preferences (P3P) 1.1 Working Group
- **Leadership Roles**
  Legislative Concerns Chair, Board of Directors
  National Association of Graduate and Professional Students (NAGPS), 2006-2008
  Vice President for External Affairs
  Carnegie Mellon Graduate Student Assembly, 2006-2008

Awards and Nominations

- **ISR Best Published Paper**, 2012

  The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study, received the Best Published Paper Award at the 2012 INFORMS Conference (with J. Tsai, L. Cranor, and A. Acquisti).
- **SOUPS Best Paper Award**, 2012

  Android Permissions: User Attention, Comprehension, and Behavior, received the Best Paper Award at the Symposium on Usable Privacy and Security (with A. P. Felt, E. Ha, A. Haney, E. Chin, and D. Wagner).
- **CHI Best Paper Nominee**, 2011

  Of Passwords and People: Measuring the Effect of Password-Composition Policies, received an honorable mention at CHI 2011 (with with S. Komanduri, R. Shay, P. G. Kelley, M. Mazurek, L. Bauer, N. Christin, and L. F. Cranor).
- **CHI Best Paper Nominee**, 2008

  You've Been Warned: An Empirical Study on the Effectiveness of Web Browser Phishing Warnings, received an honorable mention at CHI 2008 (with L. Cranor and J. Hong).
- **Tor Graphical User Interface Design Competition**, 2006

  Phase 1 Overall Winner (with L. Cranor, J. Hong, P. Kumaraguru, C. Kuo, S. Romanosky, J. Tsai, and K. Vaniea).
- **University of Virginia Dean's List of Scholars**

  I was included on the Spring 2003 and 2004 Dean's List of Scholars.
- **Publisher's Clearing House Finalist**

  I *may* already be a winner.

Last modified February 2013.