Joshua Tan

jstan@cs.cmu.edu | https://joshktan.com | jtan189 | joshua-s-tan

Summary

Research scientist with a security background and a software-engineering foundation. Experience in all phases of usability research including experimental design, application development, data analysis, and publication of results. Highly proficient at quantitative data analysis using both statistical and machine-learning approaches.

Education

Carnegie Mellon University (School of Computer Science) *Pittsburgh, PA*

Ph.D. in Societal Computing *Aug 2014–Sep 2020*

- Advisors: Lorrie Faith Cranor, Lujo Bauer
- Thesis: Practical security guidance for authentication-system designers
- Thesis committee: Matt Fredrikson (CMU), Mary Ellen Zurko (MIT Lincoln Laboratory)

M.S. in Societal Computing *Aug 2014–May 2019*

North Dakota State University *Fargo, ND*

B.S. in Computer Science (GPA: 3.97/4.00) *Aug 2009–May 2014*

Employment

Carnegie Mellon University *Pittsburgh, PA*

Postdoctoral Researcher, CyLab Security and Privacy Institute *Oct 2020–Present*

Led research study on 2FA, including quantitative and qualitative data analysis.

- Designed survey protocol to investigate reasons inhibiting voluntary adoption of two-factor authentication (2FA)
- Collaborated with Information Security Office to collect and analyze data on perceptions of 2FA security and usability

Carnegie Mellon University *Pittsburgh, PA*

Graduate Research Assistant, CyLab Security and Privacy Institute *Aug 2014–Sep 2020*

Performed multi-disciplinary research combining ideas from computer security, privacy, human-computer interaction, and psychology. Employed user studies to understand effective security and privacy in realistic, user-adverse settings. Analyzed data using statistical analysis and machine-learning modeling approaches.

- Designed and evaluated password policies incorporating a neural-network-driven minimum-strength requirement
- Used statistical techniques (hypothesis testing, mixed models, model selection, survival analysis) to analyze experimental user-study data and recommend practical, effective security/privacy designs
- Applied machine-learning techniques to text passwords, including: strength classification using logistic regression (Scikit-Learn) and unsupervised representation learning using LSTM autoencoders (Keras, TensorFlow)
- Performed fuzzy clustering to latent representation of passwords to learn high-level patterns and used t-SNE projections to visualize latent space in 2D
- Applied model-explanation techniques such as LIME and Integrated Gradients to explain clustering predictions
- Cleaned, transformed, and visualized data using R (tidyverse, lattice) and Python (NumPy, Matplotlib, pandas, Seaborn)
- Implemented interactive, Web-based, user-study activity (Django, JavaScript) to evaluate effective security of textual and graphical key-fingerprint representations in realistic, adverse settings
- Designed and implemented experimental study to examine the effects of scenario realism and purpose of use on users' privacy valuations for actual personal data (Google People API, Qualtrics, Django)

MIT Lincoln Laboratory *Lexington, MA*

Research Intern (Mentor: Richard Shay) *May 2018–Aug 2018*

Used human-subject research methods to perform a case study on a deployed security system.

- Collaborated with internal security departments to conduct a human-subjects user study on smart-card-based 2FA
- Tasks included survey protocol design, staff interviews, data analysis, and write-up of results in conference-submitted paper

Thomson Reuters *Eagan, MN*

Software Engineer Intern *May 2014–Aug 2014*

Worked on tasks related to Hadoop job scheduling and Linked-Data applications.

North Dakota State University *Fargo, ND*

Research Assistant, McNair Scholar (Mentors: Hyunsook Do, Brian Slator) *Aug 2012–May 2014*

Undergraduate research assistant on projects related to software testing and code optimization.

- Developed execution-path constraint collector and solver in Java for efficient, context-sensitive regression testing on large PHP applications
- Assisted in development of interactive, multi-user, virtual learning environment written in Java

University of California, Berkeley *Berkeley, CA*

Research Intern, Team for Research in Ubiquitous Secure Technology (TRUST) *June 2013–Aug 2013*

Led undergraduate research team on security project investigating developer adoption and user understanding of iOS permission requests. Resulted in a first-author publication at a top-tier venue.

- Developed Python toolchain to automate decryption and analysis of 4,400 iOS apps on jailbroken iPhone
- Performed static analysis to quantify adoption of purpose strings in permission requests to access-controlled resources

Publications

Practical recommendations for stronger, more usable passwords combining minimum-strength, minimum-length, and blocklist requirements
Joshua Tan, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor
*Proceedings of the 27th ACM SIGSAC Conference on Computer and Communications Security (**CCS '20**)*, 2020

Comparing hypothetical and realistic privacy valuations
Joshua Tan, Mahmood Sharif, Sruti Bhagavatula, Matthias Beckerle, Michelle L. Mazurek, Lujo Bauer
*2018 Workshop on Privacy in the Electronic Society (**WPES '18**)*, 2018

Self-driving cars and data collection: Privacy perceptions of networked autonomous vehicles
Cara Bloom, **Joshua Tan**, Javed Ramjohn, Lujo Bauer
*Thirteenth Symposium On Usable Privacy and Security (**SOUPS '17**)*, 2017

Can unicorns help users compare crypto key fingerprints?
Joshua Tan, Lujo Bauer, Joseph Bonneau, Lorrie Faith Cranor, Jeremy Thomas, Blase Ur
*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (**CHI '17**)*, 2017

(Do not) Track me sometimes: Users' contextual preferences for web tracking
William Melicher, Mahmood Sharif, **Joshua Tan**, Lujo Bauer, Mihai Christodorescu, Pedro Giovanni Leon
*Proceedings on Privacy Enhancing Technologies (**PETS '16**)*, 2016

The effect of developer-specified explanations for permission requests on smartphone user behavior
Joshua Tan, Khanh Nguyen, Michael Theodorides, Heidi Negrón-Arroyo, Christopher Thompson, Serge Egelman, David Wagner
*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (**CHI '14**)*, 2014

Professional Service & Teaching Experience

2018–2021	External reviewer for ACM Transactions on Privacy and Security (TOPS)
2019	External reviewer for IEEE Transactions on Dependable and Secure Computing
Spring 2017	TA for 18-732: Secure Software Systems
Fall 2015	TA for 15-421: Information Security & Privacy

Skills

Programming/Tools	Python, R, Java, SQL, JavaScript, Linux, Git, Docker, LaTeX
Libraries/Frameworks	dplyr, ggplot2, Keras, TensorFlow, Scikit-Learn, pandas, NumPy, Matplotlib, Seaborn, Django
Methods	Statistical analysis, Machine learning, Data visualization, Experimental design, User studies