Camelia Simoiu

csimoiu@stanford.edu
Tel. (650) 656–3420
https://web.stanford.edu/~csimoiu/
https://github.com/camioux

Education

PhD, Computational Social Science, **Stanford University**, USA 2014 – Present

MS, Artificial Intelligence, **University of Amsterdam**, Netherlands 2013 – 2014

BS (with distinction), **University of Toronto**, Canada 2005 – 2009
Double major in Applied Statistics and Economics

Skills

- **Proficient in**: Python, MySQL, R, Hive, (Py)Spark, Unix, Stan, Vowpal Wabbit.
- Excellent written and verbal communication skills. Confident and professional speaking abilities.
- Deep knowledge of supervised and unsupervised machine learning algorithms.
- Detail-oriented, passionate about telling stories with data, and making data-driven decisions.

Employment

PhD Candidate, **Stanford University**, CA 2014-Present

- 5+ years of experience with end-to-end mixed methods research projects using modern computational and statistical tools (e.g., machine learning, online experiments, Bayesian inference).
- Develop a risk assessment tool to predict likelihood of ransomware infection for consumers.
- Scan IPv4 space to extract TLS settings of web servers using ZMap; use agglomerative clustering and network analysis to understand the reasons for server misconfigurations.
- One of two lead researchers on the Stanford Open Policing project. Set up a pipeline to clean, process, and run statistical models on unstructured data from over 150+ million traffic stops across the U.S., in order to assess disparities in police interactions with the public.
- Develop hierarchical latent variable Bayesian model to quantify bias in decision-making. Work presented at TedX, publication in Nature Human Behavior *Vol. 4, 2020.*

Researcher, **Google**, Sunnyvale, CA Summer, 2019

- Conduct statistical analysis of the risk factors predisposing GMail users worldwide to be targeted with e-mail based phishing and malware attacks, under k-anonymity constraints.
- Feature engineering and selection; run exploratory analysis and visualization; fit binomial, poisson, gradient boosting models (in xgboost) via GoogleSQL, R, and Python.
- Results motivated a cross-team effort to build a new database logging phishing experience.

Researcher, **Symantec**, Mountain View, CA Summer, 2017

- Develop a supervised model to predict susceptibility to malware infection for over 5 million Norton Safe Web users; achieved >85% AUC.
- Use data mining techniques to define and extract relevant features from historical web browsing and telemetry logs.

Fellow, **Eric Schmidt Data Science for Social Good**, University of Chicago, IL Summer, 2013

- Develop a supervised-learning approach to predict future energy savings of building retrofits.
- Detect anomalous patterns of energy consumption, profile usage, and predict peaks for three years of hourly energy usage data from over 6,000 commercial buildings.

Academic papers

Working papers

Measuring the impact of defaults and online recommendations on security decisions: An empirical analysis of TLS configurations on the web
C. Simoiu, Z. Durumeric and W. Nguyen. [Under Review]

Predicting the risk of malware from web browsing behavior
C. Simoiu, C. Gates, S. Goel. [Under Review]

Peer reviewed publications

Determinants of phishing targeting and susceptibility: A large-scale study of Gmail user experiences
C. Simoiu, K. Thomas, E. Bursztein, A. Zand., *ACM Internet Measurement Conference (IMC)*, 2020.

"I was told to buy a software or lose my computer. I ignored it": A study of ransomware
C. Simoiu, C. Gates, J. Bonneau, S. Goel., *Symposium on Usable Security and Privacy* (*SOUPS)*, 2019.

Developing data science tools for improving enterprise cyber-security
Alan Turing Institute, 2018. [Technical Report]

Quantifying Systemic Cyber Risk, Report on the "Connectedness in Cyber Risk" Workshop
*World Economic Forum's Global CRQ Network***,** 2018. [Technical Report]

A large-scale analysis of racial disparities in police stops across the United States
E. Pierson, **C. Simoiu**, J. Overgoor, S. Corbett-Davies, D. Jenson, A. Shoemaker, V. Ramachandran, P. Barghouty, C. Phillips, R. Shroff, and S. Goel. *Nature Human Behaviour, 1-10*, 2020.
[**Media coverage:** The Economist, The New York Times, National Geographic, The Los Angeles Times, The Washington Post, The Daily Show, CNN, and NBC News.]
[**Taught at**: Columbia, Caltech, MIT, Princeton, and Stanford]

Studying the "Wisdom of Crowds" at Scale
C. Simoiu, C. Sumanth, A. Myore, S. Goel., AAI Conference on Human Computation and Crowdsourcing *(HCOMP)*, 2019. [**Best Paper Award]**

The Problem of Infra-marginality in Outcome Tests for Discrimination
C. Simoiu, S. Corbett-Davies, S. Goel., *Annals of Applied Statistics, Vol 11-3,* 2017.
[TedX talk: here **] [** Nominated for ASA **Outstanding Application award]**

Crowd Research: Open and Scalable University Laboratories
R. Vaish, S. Gaikwad, G. Kovacs, **C. Simoiu,** M. Bernstein, et. al. , *ACM Symposium on User Interface Software and Technology (UIST), 2017.*, 2017. **[Best Paper Honorable Mention Award]**

Service

I have served on the program committee for the following conferences:

- Reviewer, RSA Conference 2017-2020
- Reviewer, Journal of Quantitative Criminology 2020
- Reviewer, International Conference on Machine Learning (ICML) 2019
- Reviewer, International Conference on Learning Representations (ICLR) 2019
- Local chair, ACM Conference on Online Social Networks (COSN) 2015

Teaching & Leadership

Co-Instructor, **Stanford Universiy** Winter 2018

- Initiated and co-led a new semniar focused on behavioral cyber security at Stanford, *"Rethinking cybersecurity Behavioral and economic perspectives"*, in collaboration with Professor Dan Boneh and the Stanford Cyber Initiative.

Teaching Assistant, **Stanford University**

- Applied Statistics (MS&E125), TA for Sharad Goel Winter 2018
- Social Networks: Theory and Methods (MS&E382), TA for Noshir Contractor Winter 2016

Founder, **Data Analysis Workshop Series**, University of Toronto, ON 2009 – 2012

- Developed and taught wo 10-week courses for undergraduate students on data analysis techniques and Visual Basics for Applications (VBA) programming (attendance approximately 50 students).
- Create and develop course material, lectures, assignments, organize guest speakers.
- Workshop subsequently integrated into the actuarial science program at the University of Toronto.

Selected Honors [most recent]

Grace Hopper Fellowship, Twitter (declined)	2019
Best Paper Award, HCOMP 2019	2019
Cyber Security Scholarship, ISAC	2018
StanCon Scholarship, Stan Developement Team	2017
ACCEL Fellowship, Stanford University	2016-07
Winner of the Knight News Challenge Fellowship Challenge	2015
RSA Security Scholar	2015
Stanford Graduate School of Engineering Fellowship	2014
Data Science for Social Good Fellowship, University of Chicago	2013
International Exchange Bursary, University of Toronto	2007
Blue-Lands Scholarship, University of Toronto	2006
Victoria College Regents Scholarship, University of Toronto	2006
School Champion, Cayley Contest, Canadian Mathematics Competition	2003
Placed in top 5% nationally on Pascal & Fermat Contests, Canadian Mathematics Competition	2002-04

Personal

Languages: Fluent in English and Romanian (native). Conversant in French.

In my spare time, you'll find me exploring California's back roads on a road bike, waking up at 3am to climb mountains in the snow, diving, and attempting to produce tolerable sounds on the cello.