

Piotr Sapiezynski, PhD

+1 650 646 8163 sapiezynski@gmail.com Somerville, MA

Profile

I am a data scientist with 6+ years of experience in international, interdisciplinary teams. I use statistics and machine learning to derive insights from data about behaviors of people and systems, analyze networks, and audit black-box algorithms with an emphasis on privacy, transparency, and fairness.

Experience

POSTDOCTORAL RESEARCHER, NORTHEASTERN UNIVERSITY; BOSTON, MA, USA – 2017-PRESENT

I work on fairness in machine learning and privacy with Prof. Alan Mislove and Prof. Christo Wilson. Our research on Facebook shadow profiles has been covered in BBC, Gizmodo, TechCrunch, and others.

POSTDOCTORAL RESEARCHER, TECHNICAL UNIVERSITY OF DENMARK – 2016-2017

I advised graduate students' theses on mobility and network analysis. Our work on mobility constraints was covered in The Economist, Scientific American, and others. Our work on social bots was covered in Mother Jones and EurekAlert!, and others.

PHD CANDIDATE, TECHNICAL UNIVERSITY OF DENMARK – 2013-2016

I had major contributions in designing, building, and deploying a large scale computational social science study - the Copenhagen Networks Study. We collected multi-channel, high resolution (multiple data samples each minute per user over the course of two years) behavioral and network timeseries data from 1000 students using smartphones, online questionnaires, and APIs to online social networks. Our work on Android privacy was covered in The Atlantic, Computer World, and others, as well as referred to in related court cases and patents. Our work on deploying computational social science experiments was covered in Forbes, New Scientist, Die Zeit, and others.

INDEPENDENT DATA COLLECTION CONSULTANT – 2015-2016

I designed and implemented an online surveying platform which was then used by the researchers at Copenhagen University to study the effects of computer games on children.

SCIENTIFIC ADVISOR, YOU TECHNOLOGY – 2014-2015

I implemented my findings from the doctoral work on privacy and data analysis in a commercial context. You Technology was a startup offering a solution for data storage, analysis and access based on the Personal Data Store idea.

RESEARCH ASSISTANT, TECHNICAL UNIVERSITY OF DENMARK – 2011-2013

I developed mobile apps for Android (Java) and Symbian (C++) for research projects on epilepsy (reading EEG signals) and social networks.

STUDENT TEST ENGINEER, OTICON A/S – 2011-2013

I developed Android apps for personalized hearing aids. I created a test automation framework interfacing proprietary hardware, smartphones, and feature phones (MATLAB). I wrote control software for a robot used in auditory tests as well as used signal processing to report the test results.

Education

Technical University of Denmark – PhD in Network Science, 2016

Technical University of Denmark – MSc in Digital Media Engineering (w/ Honors), 2013

Warsaw University of Technology – BSc in Radiocommunication and Multimedia, 2010

Skills

OBTAINING DATA

Collecting smartphone data using Android's Java APIs, obtaining web data through documented APIs (e.g. Twitter and Facebook), undocumented APIs (e.g. Google protobuf, FlightTracker), and web scraping (using curl, phantomjs, and Selenium).

DATA ANALYSIS

Machine learning with Python (scikit-learn), data wrangling with Python (numpy, scipy, pandas, multiprocessing), bash, SQL, MongoDB. I know how to effectively employ remote machines and cloud platforms, including Amazon AWS.

DATA VISUALIZATION AND PRESENTATION

Python (matplotlib, networkx), Latex, Keynote, Adobe Illustrator, HTML, JS, CSS, PHP.

Selected publications and preprints

See full list at <https://scholar.google.com/citations?user=3UZwE6sAAAAJ>

- *Evidence for a conserved quantity in human mobility*, L Alessandretti, P Sapiezynski, V Sekara, S Lehmann, A Baronchelli, Nature Human Behavior, (2018)
- *Investigating sources of PII used in Facebook's targeted advertising*, G Venkatadri, E Lucherini, P Sapiezynski, A Mislove, PoPETS, (2018)
- *Inferring Person-to-person Proximity Using WiFi Signals*, P Sapiezynski, A Stopczynski, DK Wind, J Leskovec, S Lehmann, Proc. ACM IMWUT, (2017)
- *Evidence of Complex Contagion of Information in Social Media: An Experiment Using Twitter Bots*, B Monsted, P Sapiezynski, E Ferrara, S Lehmann, PLOS ONE, (2017)
- *Opportunities and Challenges in Crowdsourced Wardriving*, P Sapiezynski, R Gatej, A Mislove, S Lehmann, Proceedings of the 2015 ACM Conference on Internet Measurement Conference, (2015)
- *Tracking Human Mobility using WiFi signals*, P Sapiezynski, A Stopczynski, R Gatej, S Lehmann, PLOS ONE, (2015)
- *Measuring large-scale social networks with high resolution*, A Stopczynski, V Sekara, P Sapiezynski, A Cuttone, MM Madsen, JE Larsen, S Lehmann, PLOS ONE, (2014)
- *Measuring personalization of web search*, A Hannak, P Sapiezynski, AM Kakhki, B Krishnamurthy, D Lazer, A Mislove, C Wilson, WWW (2013)

Service

I have reviewed for Nature Scientific Data, PLOS ONE, EPJ Data Science, WWW '14 & '16, SocInfo '14, ICWSM '15, WebSci '15, CompleNet '17 & '18, FAT* Conference '18, CSCW '18. I have served as a program committee member for ICWSM '16 - '19.