

Talha OZ

Data Scientist • tozcss@gmail.com • San Jose, CA • 650-644-7138 • talhaoz.com

Summary

Leading the analytics research team in developing new technologies, metrics, algorithms, and models. Delivering actionable insights to stakeholders with 10+ years of experience in computational and data sciences. Guru in Python scientific computing stack.

Technical Skills

- *Python scientific stack*: Pandas, NumPy, Matplotlib, StatsModels, NetworkX, scikit-learn, Jupyter
- *Data modeling and retrieval*: SQL, JSON, BeautifulSoup, Selenium, NLTK, MongoDB, Postgres
- *Visualization and spatial analysis*: Seaborn, Plotly, GeoPandas, GIS, Shapely, GeoServer, Leaflet
- *Miscellaneous*: Java, junit, MASON, R, MATLAB, Perl, Flask, CGI, PHP, Linux, SSH, Docker, JIRA

Professional Experience

Research Area Lead - Humanyze, Menlo Park, CA Oct 2019 – Present

Computational Social Scientist / Data Scientist - Humanyze, Palo Alto Oct 2017 – Oct 2019

- Developing analytical logic/approach for GTM (go-to-market) strategy
 - Modeling indicators/metrics and benchmarking them
 - E.g., Overwork indicator, focus time indicator
 - Validating indicators/metrics with survey responses and scientific literature
 - Digital Trails of Work Stressors: <https://osf.io/wxcqp/>
- Integrating new data sources
 - Hardware: Sensors (BLE beacon cards & collectors)
 - Capturing physical location and interaction (badge)
 - Software: Workplace messaging apps (Slack & MSFT Teams)
 - Capturing time allocation and interaction
- Supporting Advanced Custom Analytics
 - Analyzed data from wearables and collaboration tools such as e-mail, calendar, and Slack
 - Methods: social network analyses, timeseries analyses, statistical significance tests, etc.
- Marketing and Sales enablement (reports)
 - Published two single-author papers [[Oz, 2020](#)], [[Oz, 2018](#)]] on People Analytics
 - Gave two lectures at Stanford University (an MBA class and a sociology class)
 - Preparing internal reports with Product, GCO and marketing teams
 - Presenting research area work to the rest of the company
- Collaborating with researchers and supervising interns
 - Project: Predicting (pre)mature promotion using Network Embeddings
 - Avi Goyal (Stanford), [Amir Goldberg](#) (GSB) and [Sameer Srivastava](#) (Haas)
 - [Project](#): The effects of temporal distance on communication behaviors
 - Tommy Fang (HBS), Jasmina Chauvin (Georgetown) and Raj Choudhury (HBS)
 - Project: Day-to-day communication analyses with Graphlets and Sequence analysis
 - Ryan Compton (PhD advisor [Steve Whittaker](#))

Researcher - Center for Social Complexity, Fairfax, VA Jun 2016 – Aug 2017

- Synthesized social network relations in a population using U.S. Census data and generated road networks and living spaces in a spatially explicit agent-based simulation ([Notebook](#), [Paper](#))

Researcher – Machine Learning and Inference Lab, Fairfax, VA Oct 2011 – Aug 2016

- Developed a [simulator](#) using non-Darwinian (rule-learning) genetic algorithms
- Taught machine learning and web programming courses (WEKA, SQL, CGI, Python)

Talha OZ

Data Scientist • tozcss@gmail.com • San Jose, CA • 650-644-7138 • talhaoz.com

- Implemented a patented machine learning model to classify ADLs from progress notes (Java)
- Consulted the faculty and researchers on data management and analysis practices
- Managed virtualization, security, database, mail, web, SVN, and Dropbox-like servers

Research and Teaching Assistant –University of Nevada, Reno, NV

Aug 2008 – Dec 2010

- Assisted teaching Data Structures, Programming Languages Concepts, Principles of Operating Systems, Computer Network Systems, Computer Security, and Calculus II courses
- Developed the probing module of a distributed Internet topology mapping system ([Thesis](#), [Paper](#))

Wireless Technologies Intern – Bosch Research, Palo Alto, CA

May 2009 - Aug 2009

- Implemented the required modules for a 6LoWPAN based web server demonstrator to achieve the idea of IoT by integrating tiny sensors into TCP/IP networks

Education

PhD in Computational Social Science (A.B.D.) – GMU

Jan 2014 – Present

MS in Computer Science (A.I. and Data Mining) – GMU

Aug 2011 – Dec 2013

MS in Computer Science (Complex Networks & Internet) – UNR

Aug 2008 – Dec 2010

BS in Computer Science (Wireless & Satellite Networks) – BOUN

Sep 2003 – Jun 2008

Personal Projects

- Showed how to draw challenging inferences from “big social data” about the source for blame; the partisan predisposition; the concerned geographies; and the contagion of complaining ([code](#), [paper](#))
- Showed the extent of online agenda building efforts of U.S. Congress members by predicting their political parties with 95% accuracy solely by examining the topics they tweet about ([code & paper](#))
- Created a demo geo-web application integrating new features into AirBnB ([code and demo video](#))
- Measured, analyzed, and visualized geo-popularities of US Governors on Twitter ([code & paper](#))
- Analyzed the news sentiment and public gate keeping behaviors on Twitter ([paper & code](#))
- Extending on selective exposure theory, proposed an audience similarity metric for news organizations and introduced methods for measuring media-party parallelism ([code & paper](#))
- Created a web-browser extension that facilitates getting Twitter data into Excel: twlets.com from idea to product: design, engineering (back- & front-end), data analysis, and monetization

Talks, Grants, Services, and Scholarships

- Program Committee Member at [SBP-BRIMS](#) (Conference on Social Computing, Behavioral-Cultural Modeling & Prediction and Behavior Representation in Modeling and Simulation) (2019, 2020)
- Presented [my work](#) in the Conference on Information and Knowledge Management (CIKM) (2016)
- Received travel grant and presented [my work](#) in PolNet8 Conference in Portland, OR (2015)
- Received Political Networks Fellowship Award and presented [my work](#) in Montreal, QC (2014)
- Received honorable mention (\$1,000) in DC Tech Hackathon by D&B in Chevy Chase, MD (2013)
- Received best poster scholarship award on machine learning and health informatics (2012)
- Received gold medal in a Data Mining [Hackathon](#) by ACM SF Bay Area Chapter on Kaggle (2012)

Selected Coursework

- | | | |
|--|----------------------------------|--|
| • Computational Analysis of Social Complexity | • Geographic Information System | • Data Mining |
| • Spatial Agent-based Models of Human-Environment Interactions | • Geospatial Intelligence | • Algorithm Analysis |
| • Complexity Theory in Social Sciences | • Web-based GIS | • Social Networks |
| • Geosocial Media | • Big Data | • Artificial Intelligence |
| | • Data Mining on Multimedia Data | • Semantic Web & Knowledge Engineering |

Talha OZ

Data Scientist • tozcss@gmail.com • San Jose, CA • 650-644-7138 • talhaoz.com

- Software Modeling
- OO Software Specification & Construction
- Complex Networks
- Computer Networks
- Internet Protocols
- Genetic Algorithms
- Combinatorics & Graph Theory
- Wireless Networks
- Satellite Networks