

Matt Mollison, Ph.D.

COFOUNDER & CHIEF DATA SCIENTIST AT ANSARO

San Francisco, CA, USA

✉ matt.mollison@gmail.com | 🏠 mattmollison.com | 📄 github.com/warmlogic | 🔗 [linkedin.com/in/mattmollison](https://www.linkedin.com/in/mattmollison) | 🌐 www.ansaro.com

Summary

Data scientist and machine learning enthusiast. Cognitive neuroscience Ph.D. rooted in experimental psychology. Driven to do impactful work, especially in areas that benefit people and society. Values learning, communication, and planning to enable effective teamwork.

Skills

Programming Languages	Python, SQL, R, MATLAB, Objective-C*, C* (* <i>prior experience</i>)
Analysis Techniques	Statistics, machine learning (supervised & unsupervised, classification & regression), natural language processing, neural networks, time series analysis, experimental design, data visualization, data munging
Tools & Technologies	Pandas, scikit-learn, statsmodels, Jupyter notebooks (strong opinion on when to use them), Git, AWS, GCP, Flask, fast.ai, TensorFlow, PySpark

Experience

Ansaro

San Francisco, CA

COFOUNDER & CHIEF DATA SCIENTIST

Jun. 2017 – Present

- Helped define Ansaro's cultural foundation by establishing mission, values, and feedback/reflection sessions from the beginning.
- Led all aspects of data science-related efforts including scoping projects, dev environment automation, analysis, model training/versioning, and communicating results.
- Developed the content for Ansaro's structured interview platform, nearly 200 interview questions across 16 topic areas, based on academic research related to job performance predictive validity.
- Created the web app POC of Ansaro's *AI Notetaker* (interview audio recording and transcription). Conducted user research to design this feature. Our team then developed this in the Ansaro platform, enabling collection of valuable interview content.
- Developed and delivered interpretable machine learning models for Teach For America (an Ansaro customer) to predict application withdrawal risk based on applicant information and written essays.
- Built and ran the pipeline to parse information from 800,000+ resumes into structured data for Dropbox (an Ansaro customer), standardize the parsed data, and upload to Greenhouse, to understand hiring patterns and drivers of quality of hire.
- Planned and executed a project for a Harvard Business School professor that explored and analyzed a large text corpus using topic modeling and classifiers to surface temporal trends of how women are treated in the workplace (research ongoing).
- Contributed to the company blog, communicating ideas that are important to our company, particularly related to people analytics and data science.

Freelance Consultant

San Francisco, CA

DATA SCIENTIST

Apr. 2017 – May 2017

- Strategic planning for the analysis and processing/machine learning pipeline for a new sports-related mobile application.
- Analyzed time series motion sensor data collected with the mobile phone app.
- Trained and assessed machine learning models to predict current user behavior based on motion data.
- Supervised an Insight Data Science fellow in using the pipeline and models that I developed, in their project.

Silicon Valley Data Science

Mountain View, CA

DATA SCIENTIST

Jan. 2015 – Mar. 2017

- Trained deep neural networks (RNN/LSTM) to create a speech recognition system, which was the seed R&D for a spinout company.
- Developed and deployed predictive models to production at a major global company, the results of which surfaced insights about customer behavior related to churn, purchasing patterns, fraud, and more.
- Created a process to train time series models (ARIMA) to forecast sales figures at a large energy company. The models reduced sales forecast error for the most recent quarter by 19.5% (roughly \$30MM).
- Helped develop a data strategy report at a large technology hardware manufacturing and supply chain management company. Through onsite research and interviews, the report detailed how to adapt and expand current data collection practices to provide analytics and data services across their verticals (retail, manufacturing, healthcare, transportation & logistics).
- Detected anomalous behavior in multi-terabyte system logs at a provider for the global travel and tourism industry.
- Co-developed an ARIMA time series modeling tutorial, presented at PyData SF 2016.

Insight Data Science

FELLOW

Palo Alto, CA

Sep. 2014 – Nov. 2014

- Created *This Is Happening*, a website for finding social events happening now.
- Detected events in real-time by streaming social media activity into a database and comparing current activity to baseline levels. Pockets of activity were used as seed locations for a clustering algorithm to label events.
- Summarized event topics using NLP, and made API calls to retrieve nearby social media (Twitter, Instagram).
- Deployed the front-end using Flask, Bootstrap, jQuery, JavaScript, HTML, CSS, and AWS.

Education

University of Colorado Boulder

Boulder, CO

PH.D. IN COGNITIVE NEUROSCIENCE

2015

- **Dissertation:** *Distributed practice and distributed representations: Investigating the spacing effect using EEG*
- Employed statistical and machine learning techniques on neural and behavioral data to assess hypotheses about memory formation, with implications for improving studying strategies in the real world.
- Communicated findings via publications and conference talks to greater psychological community.
- Developed **mat-mvm**, a MATLAB software package for organizing and analyzing neural data.
- Developed **expertTrain**, a MATLAB software framework for running psychology experiments.
- Programmed experiments in MATLAB and Python for studying human memory.
- Course instructor for *Introduction to Cognitive Psychology*.
- Teaching assistant for *Statistics*.

Brandeis University

Waltham, MA

B.A. IN PSYCHOLOGY

2005

- **Honors Thesis:** *Event-related potentials in humans during spatial navigation*
- Conducted and analyzed an experiment investigating neural correlates of high-level perception in a spatial navigation task, culminating in the above thesis and subsequent publication.

Honors & Awards

April 2016	2nd place team project , Kafka Stream Data Hackathon	San Francisco, CA
2009	Honorable Mention , National Science Foundation Graduate Research Fellowship	Boulder, CO
2003 – 2005	Dean's List , Brandeis University	Waltham, MA
2000	Eagle Scout , Troop 441	Scottsdale, AZ