Samuel Brotherton

CONTACT INFORMATION

Phone: (801) 927-0219
Email: sbrother@gmail.com
LinkedIn: https://www.linkedin.com/in/sbrother
Address: 7771 Buckboard Drive, Park City, UT 84098

EDUCATION

Harvard University, Cambridge, MA

B.A., Mathematics and East Asian Studies **Sep 2008 – May 2012**

Received highest honors for senior thesis analyzing over 200,000 Chinese blog posts, algorithmically detecting mutations in the Chinese language in response to censorship. Completed coursework in abstract algebra, Galois theory, topology, real and complex analysis, probability theory, and linguistics.

PROFESSIONAL EXPERIENCE

Nextdoor, San Francisco, CA / Park City, UT (Remote)

Staff Software Engineer **Jan 2023 – Present**

Tech lead on a team building a next generation ads platform from scratch while gradually migrating Nextdoor's existing ~$250M of revenue off of a legacy platform powered by Google Ads Manager. Extensive cloud infrastructure development to introduce, maintain, and monitor streaming and batch data pipelines using Apache Beam in Kotlin, Flink, Kafka, and TimescaleDB on Kubernetes/AWS.

Google, Los Angeles, CA / Park City, UT (Remote)

Software Engineer **Feb 2021 – Jan 2023**

Software engineer on an infrastructure team within Ads Privacy and Security, using natural language processing and machine learning to prevent policy-violating ad impressions. Our systems handled 2M+ queries per second and combined modern deep learning techniques with multi-tier human evaluation pipelines. Additionally, I had a 20% role building remote sensing infrastructure and machine learning models for a stealth climate "startup" within Google.

Authentic Artists, Los Angeles, CA / San Francisco, CA / Park City, UT (Remote)

Founding Engineer **Jan 2019 – Jan 2021**

Led a remote engineering team to build AI-powered audiovisual experiences, including a cloud deep learning server for song generation, a C++ constraint/transformation library for imposing additional musical style on its output, and an automated Max/Ableton session for audio generation. Coordinated with creative and executive teams to balance engineering and IP development with artistic and business goals. See `warpsound.ai` or `https://open.spotify.com/album/4ws0VXLBCNhk7tLO33xfD5` for examples.

Cairn Labs, Los Angeles, CA / Portland, OR / Park City, UT

Founder & Principal **Jan 2016 – Present**

Founded a successful boutique consultancy to design and build software for clients, with a focus on applications that integrate state-of-the-art NLP and machine learning. Managed a team of 2-6 depending on client load. Most projects involve a multi-year timespan including the sales, design, and implementation phases, followed by eventual transition to in-house resources. Our clients have cumulatively raised over $500M of venture funding using our prototyping and fractional CTO services. A small selection of notable projects includes:

- **i.am+**: A deep learning based conversational UI framework to power will.i.am's wireless earphones and other applications. Backed by Tensorflow, Prolog, and other technologies. Supports multiple languages, extensible dialogue flows, and a custom knowledge base. See `https://goo.gl/iMvVTt`.
- **Blue Cross Blue Shield**: A highly available machine learning inference server used to predict asthma risk in real time from claims records.
- **Hikma Health**: An offline-first, mobile first Electronic Health Records system designed for the unique challenges of mobile refugee health clinics. Currently deployed in Lebanon and Nicaragua (see Brotherton et al in *Selected Publications*).
- **Motion Insurance** A risk scoring server for car insurance, collecting realtime driving data and assigning ML-based driver risk scores at a rate of 3000+ qps (see `motioninsurance.com`).

Google, Venice, CA

Software Engineer **Jun 2014 – Mar 2016**

Worked on a small team using natural language processing and other machine learning techniques to improve advertisement quality. Led a 20% project related to mining semantic information from web data, which was adopted by several teams across different product areas. Built a named entity recognition system in C++ and a link detection algorithm that runs on very large graphs; contributed to a topic model for clustering semantic entities.

Whisper, Venice, CA

Software Engineer and Data Scientist **Mar 2013 – Apr 2014**

Sole data scientist at a rapidly expanding social media startup seeing upwards of three billion monthly pageviews. Designed and built an NLP service to extract topics and tags from posts, predict image searchterms from unstructured text, and target content to users. Implemented a new geographic search system using PostGIS that decreased search time by 90%. Worked closely with the front and backend development teams, writing production code in Erlang and Python.

Selected Publications

Brotherton, T., Brotherton, S., Ashworth, H., Kadambi, A., Ebrahim, H. and Ebrahim, S., 2022. Development of an Offline, Open-Source, Electronic Health Record System for Refugee Care. *Frontiers in Digital Health*, 4.

Demiris, G., Oliver, D.P., Washington, K.T., Chadwick, C., Voigt, J.D., Brotherton, S. and Naylor, M.D., 2022. Examining spoken words and acoustic features of therapy sessions to understand family caregivers' anxiety and quality of life. *International Journal of Medical Informatics*, 160, p.104716.

Demiris, G., Corey Magan, K.L., Parker Oliver, D., Washington, K.T., Chadwick, C., Voigt, J.D., Brotherton, S. and Naylor, M.D., 2020. Spoken words as biomarkers: using machine learning to gain insight into communication as a predictor of anxiety. *Journal of the American Medical Informatics Association*, 27(6), pp.929-933.

Stephens, M., Bensink, M., Brotherton, S., Chandler, D., Garcia, J. and Hollenbeak, C., 2016. Geographic Access to Oncology Services in the United States (US): Travel Disparities May Affect Granulocyte-Colony Stimulating Factor (G-CSF) Administration. *Blood*, 128(22), p.5905.

Stephens, M., Brotherton, S., Dunning, S., Emerson, L., Gilbertson, D., Harrison, D.J., Kochevar, J., McClellan, A., McClellan, W., Wan, S. and Gitlin, M., 2011. The End-stage Renal Disease (ESRD) Prospective Payment System (PPS) and Access to Care: Incremental Distance Traveled by Displaced Patients.

Programming Experience

Languages: Python, Kotlin, C++, Erlang, Elixir, Mathematica, C#, F#, Bash, C, LaTeX
Server Technology: Kubernetes, Docker, GCP, AWS, Azure, Cassandra, Redis, PostgreSQL, ElasticSearch, Apache Beam, Apache Flink, TimescaleDB