

Brian Matejek

COMPUTER SCIENCE · ARTIFICIAL INTELLIGENCE · DATA SCIENCE · MACHINE LEARNING · ALGORITHMS

✉ bmatejek@gmail.com | 🌐 brianmatejek.com | 📧 bmatejek | 📄 brian-matejek | 🐦 @brianmatejek

Summary

Advanced computer scientist with 10+ years of experience developing AI/ML solutions for defense, intelligence, neuroscience, and health-care communities including as Principal Investigator on a DoD program with an active Top Secret clearance. Specializes in high-assurance machine learning, uncertainty quantification, and energy-efficient AI with recent work spanning cybersecurity and clinical settings.

Work Experience

SRI International

ADVANCED COMPUTER SCIENTIST

Feb 2022 - Present

- Led a multi-year, multi-million dollar Controlled Unclassified Information (CUI) Department of Defense program as Principal Investigator (PI), working together with multiple external teams for a government stakeholder resulting in publications in a top machine learning journal (TMLR) and cybersecurity workshop (AICS at AAAI).
- Extensively employed foundation models to address emerging threats in cybersecurity. Trained language models from scratch to identify malicious network activity at a packet-level. Deployed generative models to monitor (and approve) the activity of the language models.
- Conducted research across cybersecurity, healthcare, and AI/ML — producing 17 peer-reviewed publications (3 journal, 5 conference, 9 workshop) with additional work under review.
- Worked on five major government programs from DARPA (QuICC), IARPA (MICrONS, TrojAI), ARPA-H (DIGIHEALS), and DoD funding sources. Mentored five students for these efforts culminating in publications (and open-source code) at top AI conferences and workshops.

Freelance Consulting

DATA ANALYST

Dec 2018 - Feb 2022

Client: mid-sized electrical distribution company.

- Developed a “pricing matrix” that recommends to salespeople target margins for each transaction based on their region of the country, the product class, and the customer type.
- Created a web-based interface with a Django backend and PostgreSQL database that enables the company to analyze millions of transactions over the previous eight years and identify revenue and margin trends by product, customer, and salesperson.
- Participated in weekly meetings over multiple years with the company president and others to devise an incentive program that encourages salespeople to adhere to the new pricing matrix.

Client: early-stage military-focused start up

- Connected the company’s existing desktop technology to a new web interface and added features enabling users to upload gigabyte LIDAR scans to an Amazon S3 bucket. On file upload, a lambda function converts the file into a format suitable for web-based visualization.
- Helped draft a 30+ page provisional patent for the company that included more than a dozen figures and schematics of the end-to-end pipeline. Summarized dozens of existing patents in the same domain to help differentiate the proposed technology from the prior art.
- Engaged in 3-4 hour long meetings per week with potential customers and collaborators in the Air Force, Navy, Coast Guard, and private sector. Led several meetings with research scientists in the defense sector to discuss the current state and future vision of the technology.
- Co-led a demonstration of the company’s visualization software for three potential clients on a WWII-era museum battleship.

Harvard University

DOCTORAL RESEARCHER

Aug 2016 - May 2021

- Trained a 3D convolutional neural network to learn neuronal shapes and correct errors in reconstructions of brain tissue. Devised novel, domain-guided techniques to increase throughput to gigavoxel image datasets (work published in *CVPR 2019*).
- Designed the *Compresso* algorithm to compress terabyte image dataset, achieving compression ratios exceeding $700\times$, an $8\times$ improvement over existing state-of-the-art. Implemented the solution in C++, Javascript, and Python to encourage widespread adoption in the community, processing over 100-200 megabytes of data per second (work published in *MICCAI 2017*).
- Devised algorithms so that they can be deployed in parallel across a compute cluster enabling computation on terabyte image datasets. Learned custom heuristics for smarter load balancing. Deployed the solutions using SLURM scheduling across a computer cluster enabling several years of computation time in two weeks (work published in *MICCAI 2019*, *TMI 2022*, *Nature Scientific Reports 2022*).
- Mentored a dozen students on research and engineering projects including three master’s theses. Worked with these students to develop ideas and implement solutions.

Education

Harvard University

Cambridge, MA 02138

PH.D. IN COMPUTER SCIENCE

Aug 2016 - May 2021

Princeton University

Princeton, NJ 08544

M.S.E. IN COMPUTER SCIENCE

Sep 2014 - May 2016

B.S.E. IN COMPUTER SCIENCE

Sep 2010 - Jun 2014

Skills

Programming Python, C++, C, LaTeX, PyTorch, Huggingface, OpenHands, Cython, Numpy, Pandas, Django, Sass, jQuery, Linux, SLURM
Clearance Top Secret Clearance