Tucson, Arizona
jcallahan4@gmail.com
801-473-5991

Jake Callahan

linkedin.com/in/jake-callahan
github.com/jcallahan4

Applied ML researcher with professional experience developing AI/ML models using supervised and unsupervised learning techniques. Experience deploying production systems generating $1M+ revenue. Proven experience developing Bayesian inference frameworks for national security applications at Sandia National Labs. Expert in probability/statistics, uncertainty quantification, and high-performance computing with Python and JAX. Transitioning from PhD research to industry machine learning roles.

Work Experience

Graduate Research Scientist **Sandia National Labs & The University of Arizona** **May 2021–Present**
Tucson, Arizona

- Designed a new Monte Carlo algorithm that reduced the computational cost of Bayesian optimal experimental design tasks by up to 5x compared to standard methods. Research published as first author in NeurIPS 2025.
- Engineered a scalable ML framework in Python to enable previously infeasible optimization of national seismic monitoring networks, collaborating with geophysicists and engineers to advance capabilities for scientific and national security applications.
- Developed production-ready Bayesian models with neural network surrogates for anomaly detection and sensor placement, implementing GPU-accelerated uncertainty quantification that achieved a 10× speedup.

Data Scientist **OrderBoard, Inc.** **2019–2021**
Orem, Utah

- Engineered and deployed a production NLP model using embedding techniques and cosine similarity matching that automated job-candidate pairing for thousands of requisitions, forming the core ML product that generated $1M+ in annual recurring revenue.
- Built and maintained data infrastructure for the company's ML pipeline, including data cleaning, feature extraction, and performance validation.

Automation and Cognitive Services Intern **Honeywell** **Summer 2020**
Charlotte, NC (Remote)

- Developed a data pipeline using Python and MongoDB to ingest, store, and index OCR-scanned contract data, enabling full-text search capabilities for the legal department and reducing document retrieval time.

Education

- **Ph.D. Applied Mathematics,** The University of Arizona **2023 - Present**
- **M.S. Mathematics,** Brigham Young University **2021–2023**
- **B.S. Applied and Computational Mathematics,** Brigham Young University **2016–2020**

Technologies and Languages

- **Programming**: Python, R, SQL • ML **Frameworks**: JAX, PyTorch, Scikit-learn, NumPy, Pandas • **Statistical Methods**: Bayesian inference, MCMC/SMC, Uncertainty quantification • **Tools**: Git, HPC (SLURM), MongoDB, Docker, CI/CD

Publications

- **Reverse-Annealed Sequential Monte Carlo for Efficient Bayesian Optimal Experiment Design** (Neural Information Processing Systems, 2025)
- **Analysis and Optimization of Seismic Monitoring Networks with Bayesian Optimal Experimental Design** (Geophysical Journal International, 2025)