Zachary Dawson

M.S. CS (ML), Georgia Tech — 2024–Present

zachwdawson@gmail.com | 508 446 3322 | zachwdawson.github.io

Experience

Data Engineer (part-time), Learning Analytics Lab, Georgia Tech – Atlanta, GA Aug 2025 - Present

- Develop and maintain an Azure-based learning-analytics platform: integrate new data sources, build PySpark ETL, contribute Terraform IaC and developer-experience improvements, and support dashboards/interactive visuals.

Senior Analyst: Data Engineer, Global Atlantic Financial Group – Brighton, MA Sept 2021 - May 2024

- Migrated 200+ Control-M jobs to AWS Glue with Terraform, enabling infra-as-code deployments and reducing manual scheduling.
- Introduced the company's first ML model (cosine similarity + random forest) to detect duplicate contacts, improving accuracy by 10% vs. rules and productionizing in the pipeline.
- Built an automated test framework for database and stored procedure migrations (Python + CI), replacing manual checks and catching regressions pre-deploy.
- Integrated Salesforce Bulk API to read/write prospects, adding 5,000 new records to Salesforce and the enterprise data warehouse with auditability.
- Migrated annuity CRM data to the in-house warehouse to enable enterprise reporting at scale (700,000 policies; $40B in transactions).
- Automated commission allocation service used by sales ops to reconcile and split $5M+ in payouts.

Bioinformatics Co-op, MIT BioMicro Center – Cambridge, MA Jan 2020 – June 2020

- Improved and monitored next-gen sequencing pipelines; ensured timely delivery of 2–4 TB/week of data.

Server/SRE Co-op, Skillz Inc. – San Francisco, CA Jan 2019 – June 2019

- Created load-testing infrastructure and delivered weekly production/staging performance reports to engineering leadership.

Education

Georgia Institute of Technology MS in Computer Science | GPA: 4.0 Aug 2024 – May 2026
Relevant Coursework: Deep Learning, Machine Learning, Bayesian Statistics, Natural Language Processing
Northeastern University Honors Program BS in Computer Science GPA | 3.7 Sept 2017 – May 2021

Technical Skills

Languages: Python, SQL, Scala, Java, Bash
ML/Stats: PyTorch, scikit-learn, PyMC, Pandas, NumPy, SciPy
Data/Orchestration: Spark/PySpark, Control-M
Cloud & IaC: AWS (S3, Glue, Redshift, Lambda), Azure (Synapse, Delta, Entra), Terraform
Other: Git, Linux, MySQL, Jenkins, Jupyter

Projects

Bayesian Golf Handicapping link

- Developed a more fair golf handicapping system using Bayesian Methods to better predict future golf score distributions, achieving 7% fairness discrepancy compared to 10% for existing USGA handicapping)
- Tools Used: Python, PyMC, Pandas, NumPy, SciPy, Selenium, Matplotlib

Golf Swing Action Recognition link

- Improved upon golf swing frame labeling benchmark (81.6% vs 76.1% PCE) using improvements in Computer Vision techniques including 3D CNNs and custom loss functions.
- Tools Used: Python, PyTorch, Pandas, NumPy, Matplotlib