

Andrew Simonson

Data Scientist — Machine Learning & Analytics Engineering

asimonson1125@gmail.com — (240)-707-1486 — asimonson.com — LinkedIn — GitHub

EDUCATION

Rochester Institute of Technology

Rochester, NY

M.S. Data Science (Capstone Fall 2026) — *B.S. Computer Science*, GPA 3.63, Dean's List (Dec 2024)

Dual Cluster in Data Science and AI — Minor in International Relations — **UCSanDiegoX Data Science MicroMasters**

EMPLOYMENT

Data Scientist / Data Engineer — Ecolab

January 2024 – Present

Saint Paul, MN

- Developing surge preparation models and analytical tooling under the Data Science Technology Platforms team
- Liaison for Microsoft in 100-member AI accelerator to develop RushReady AI for the service industry
- Designed scalable ETL platform to route vendor data with high observability using Azure Data Factory and Databricks
- Built anomaly detection and recommendation models, increasing speed of service by 20%

Data Engineer — Dow Chemical

January 2023 – May 2023

Freeport, TX

- Developed automated data processing and analysis tools for Reactive Chemistry analysts, replacing 4 deprecated tools
- Designed report abstraction pipeline preserving data integrity across 9 experimental variations
- Product became the interdepartmental data standard, saving 1,020 hours (~0.65 FTE) annually

PROJECTS

Human vs AI Code Smell Study (asimonson.com)

- Analyzed 85K+ code definitions comparing AI vs human code maintainability via static analysis and non-parametric statistics
- Found AI code carries 3–4x more structural debt invisible to reviewers; recommends automated quality gates for AI-generated PRs

Generalized Trust Review (asimonson.com)

- Cross-national correlation analysis ($r=0.624$, $p<0.001$) of media and science trust across 128 countries with subgroup OLS regression
- Trust transfer attenuates 5.7x in low-press-freedom regimes; quartile progression from 12.3% to 26.2%

Antietam Watershed Monitoring

- ArcGIS choropleth dashboard monitoring stream temps and biotic indices across Upper Potomac sub-watersheds
- Flags exceedances against 90th-percentile thresholds for state environmental review

Slate (GitHub)

- Surfaces optimal meeting windows by vectorizing iCalendar data into discrete-time signals with weighted constraint scoring
- Periodic fold analysis across weeks reveals recurring availability trends for 80-member org

SKILLS

Languages: Python, SQL, R, MATLAB, JavaScript, Rust, C#, $\LaTeX$

Platforms & Tools: MS Azure, Databricks, Apache Spark, Docker, Kubernetes, ArcGIS

Methods: Machine Learning, Data Mining, PyTorch, Pandas, Scikit-Learn

ACTIVITIES

Computer Science House — E-Board & Presenter (csh.rit.edu)

Aug 2021 – Present

Rochester, NY

- Presented seminars on web scraping, analytics, GIS
- Built and maintained web services for 80+ members

CERTIFICATIONS & AWARDS

Certified by Databricks: **ML Model Deployment**, **Retrieval Agents**, **MLOps**

Certified by Microsoft: **Data Analysis** Career Essentials, **Docker** Foundations

Hackathon Awards: **CSHacks** – Best Use of AI (Paychex) — **Tiger Games** – Social Impact Winner — **HACK.COMS '24** – Accessibility Winner