Audrey Houghton (she/her)

github.com/audreymhoughton | linkedin.com/in/audreyhoughton | audreymhoughton@gmail.com
+1-612-462-1481 | Google Scholar

Technical Skills

Software

Proficient with Python, UNIX/Linux, GitHub, LaTeX, EMACS, VSCode, Singularity, Docker, and virtual environments. Familiar with Matlab, R, IDL, JavaScript, Node.js, npm, Go, Ansible, AWX, RedHat, ReadTheDocs, and Google Apps Script (via clasp). Knowledgeable in CI/CD pipelines (e.g., CircleCI, GitHub Actions) and deep learning AI (nnU-Net). Applied skills in DevOps, data engineering, and software engineering roles.

Data

Proficient in high-performance computing and data systems, including AWS (S3), Ceph, SLURM, Kubernetes, Globus, Grafana, and Prometheus. Experienced in ETL-based data warehousing, batch processing, SQL querying, and testing and validation of data pipelines. Skilled in statistical analysis, benchmarking, causal inference, predictive modeling, GIS-based spatial analysis, and electronic health record (EHR) data integration. Experienced in automation workflows including Salesforce API integration, Microsoft Teams automation, and compliant web scraping practices.

Product Management

Proficient in cross-functional and customer communication, Kanban boards, Jira, and documenting SOPs, methodologies, terminology, and codebases. Experienced with MVP development life cycles and familiar with Agile and Scrum frameworks. Applied business process automation to improve sales and reporting workflows.

Experience

University of Minnesota – Department of Psychiatry & Behavioral Sciences *Dec 2025 - Present*
Research Professional III Minneapolis, MN (Remote)

- Develop and maintain statistical and computational analysis pipelines for GIS-linked electronic health record data
- Implement causal inference, predictive modeling, and machine learning workflows in Python, R, and JavaScript
- Write and interpret results for publications and grant mechanisms, serving as a data and methods resource for an interdisciplinary research team

CoreWeave, Inc. *Apr 2024 - Sept 2024*
Fleet Operations Engineer Remote

- Provisioned and maintained thousands of NVIDIA GPU nodes in Kubernetes clusters using Ansible, AWX, Grafana, and Go, achieving onboarding timelines of 1–2 weeks with 6+ months average production uptime
- Diagnosed and resolved hardware/software issues across the node lifecycle, saving approximately $250,000 per node repaired
- Partnered with 10+ teams—including networking, hardware, operations, and support—to streamline troubleshooting, accelerate incident resolution, and document SOPs and post-mortems, improving efficiency and reliability
- Identified points in the node lifecycle where nodes were lost, implementing solutions that saved the company millions in revenue

Masonic Institute of the Developing Brain *Jan 2021 - Apr 2024*
Research Professional II → Research Professional III Minneapolis, MN (Remote)

- Developed Python wrappers leveraging AWS S3 storage and SLURM optimization to accelerate image processing up to 12x
- Engineered a containerized infant MRI segmentation application with deep learning models, achieving 600x speed-ups and +53% accuracy
- Built ETL-based pipelines to transfer and process petabytes of imaging data 10x faster than prior industry benchmarks, incorporating statistical error analysis and validation to ensure reliability and reproducibility
- Authored comprehensive documentation and curriculum, setting standards for reproducible neuroimaging data processing

Calyxt *Oct 2019 - Aug 2020*
Laboratory Assistant & Data Specialist Roseville, MN

- Developed Python scripts for comprehensive data tracking and analysis methods

Huntington Learning Center *Aug 2019 - Oct 2019*
Mathematics Instructor Plymouth, MN

Tutor Doctor *Jan 2019 - Jun 2019*
Mathematics Instructor Loveland, Colorado

Waste Not, LLC. *Est. 2018*
Data Manager and CEO

Amazon *Jul 2018 - Aug 2018*
Data Entry Associate Seattle, WA

Imagine Prep at Surprise *Jul 2017 - Jul 2018*
Mathematics Instructor Surprise, AZ

The MINVERA Project *Jan 2015 - May 2017*
Research Assistant Missoula, MT

- Developed a data reduction pipeline for time-domain imaging data written in Python

Rutgers University Astrophysics Department *May 2016 - Aug 2016*
Research Assistant New Brunswick, NJ

- Refactored IDL scripts into Python
- Generated models for galactic simulations through parameter optimization

Products

Sales Summary Project 2025
Automation Pipeline for Lead Generation & Sales Tracking [Proprietary]

- Designed a proof-of-concept pipeline integrating Salesforce logging, CSV rotation, and Microsoft Teams alerts
- Implemented compliant web scraping (robots.txt, throttling, retries) to demonstrate reliable, ethical data collection for potential enterprise use

JavaScript Excel Sheet App 2025
Google Apps Script Utility for Dynamic Links [Personal]

- Built a Google Sheets App Script to dynamically generate links, clean outdated entries, and trigger updates in real time
- Reduced personal job application tracking effort by automating link formatting and updates; version-controlled with GitHub using clasp, Node.js, and npm

CoreWeave Internal Documentation 2024
CoreWeave, Inc.

- Initiated and built company-wide documentation from the ground up, replacing scattered one-off documents with an organized, centralized Notion knowledge base covering terminology, SOPs, and team workflows

SLURM Wrappers 2021 - 2024
High-Performance Computing S3 and Disk SLURM-Compliant Wrappers github.com/DCAN-Labs/SLURM_wrappers

- Bash/Python wrappers adaptable to any SLURM-based data pipeline

ABCD Study Collection 3165 Data Processing 2021 - 2024
NIH Data Archive (NDA) collection3165.readthedocs.io

- Community dataset enabling access to petabytes of processed/unprocessed MRI data under BIDS standards

DCAN Labs Informational Guide 2022 - 2024
Minnesota Supercomputing Institute Data Engineering Documentation dcan-labs-informational-guide.readthedocs.io

- Comprehensive, continually-supported documentation of usage of the Minnesota Supercomputing Institute to process, analyze, and transfer MRI data

CABINET 2022 - 2023
Container-Linking Wrapper github.com/DCAN-Labs/CABINET

- Python-based universal container-linking wrapper employing FAIR principles

BIBSNet 2022 - 2023
Infant MRI Brain Segmentation Application github.com/DCAN-Labs/BIBSnet

- Containerized infant MRI application; achieved 600x faster processing and +53% accuracy

Data Processing Workshop 2022
Minnesota Supercomputing Institute (MSI) Data Processing Workshop tinyurl.com/MRIsyllabus

- An extensive all-levels MSI data processing workshop
- Used as a template for a University neuroscience course at the University of Minnesota

Education

University of Montana - Missoula, MT Aug 2013 - May 2017
Bachelors of Arts in Physics umt.edu/physics-astronomy

University of Minnesota Fall 2023
Data Management Series libguides.umn.edu/datamanagement/dmbootcamp

University of Minnesota 2023
Project Management Certificate Coursework (In Progress) ccaps.umn.edu/project-management-certificate
Completed modules: Foundations, Execution & Monitoring, Leadership, Team Communication