



Patrick Dwyer

Data scientist with 3+ years of experience in software engineering, machine learning, and scientific computing. Degrees in Mathematics and Computer Science from Northwestern University; Google Cloud [Certified](#) Professional Machine Learning Engineer. Built and productionized ML/NLP research code, developed large-scale data pipelines, and trained neural networks in PyTorch. Currently focused on applying computational methods to problems in biology. Live CV at patrickdwyer.com.

Education

BA Mathematics, BA Computer Science

Sep 2019 – Jun 2023

Northwestern University, Evanston, IL — Weinberg College of Arts and Sciences — 3.43/4.00 GPA

Projects

cwdmdc.patrickdwyer.com

Winter 2025 – Present

- Built a responsive dashboard for Chronic Wasting Disease sampling results on MDC's public ArcGIS API

Research Experience

Schwartz Lab: Research Assistant

Jul 2023 – Sep 2023

- Developed a supervised 3D convolutional neural network in PyTorch to track key mouse body parts in 3D from four synchronized camera views
- Extended PyTorch's `torch.autograd.Function` class to integrate 3D→2D point projection into autograd
- Implemented an extrinsic camera-calibration routine using Levenberg-Marquardt for Bundle Adjustment
- Manually labelled 8,544 ground-truth 2D points with a custom-built image-labeling program

Professional Experience

Manifold Group: Data Scientist

Oct 2023 – Present

- Designed and implemented internal chat bot timesheet application on GCP & Palantir Foundry, increasing time tracking compliance by 38%
- Productionized NLP research code, designing and architecting a scalable solution that reduced runtime from 6 hours to 30 minutes (~92% improvement), enabling faster iteration and broader production use.
- Implemented vector embedding similarity pipelines from inception to production, unlocking the ability for clients to analyze previously opaque free-form text and derive actionable insights.
- Optimized client stored procedures decreasing query time by 99.8% to enable ingestion of > 100GBs of data
- Developed and maintained data pipelines to ingest and process > 1TB of data from a client's database into a data lakehouse in Microsoft Fabric
- Led, designed, and implemented enterprise-level data solutions for a Fortune 500 client using PySpark and Microsoft Fabric
- Collaborated with a team to build terraform resources and design networking for a data lakehouse architecture on Azure
- Designed and developed anomaly detection application (Python, Docker, AWS EC2)

Websanity: Web Developer

Sep 2023 – Present

- Built and maintain JavaScript features for [Rulepop](#), an emerging rules reference platform for tabletop games