

Devon Peticolas

SOFTWARE ENGINEER

(732) 589 - 3470 | devon@peticol.as | peticol.as | [x](#)

About

I'm a software engineer who specializes in architecting scalable real-time and historical data systems, turning data science and machine learning research into production systems, and building compelling data-driven products. I believe in leading by teaching. I regularly present at industry and academic conferences and teach Data Science and Machine Learning at Columbia University's Pre-College Program.

I live in NYC and am seeking an in-office individual contributor engineering role focused on machine learning, AI, and delivering end-user impact.

Work Experience

Scale AI

New York, NY

STAFF SOFTWARE ENGINEER

Nov 2025 – Present

- Forward-deployed engineer building agentic AI solutions for enterprise customers across healthcare and education.
- Technical team lead on natural-language to SQL AI platform capabilities.

Oden Technologies

New York, NY

PRINCIPAL SOFTWARE ENGINEER, DATA SCIENCE TEAM

Oct 2017 – Oct 2025

- Founded Oden's Data Science team; built initial infrastructure, defined data standards, hired and mentored Data Scientists, and led technical product development.
- Architected and implemented scalable real-time streaming pipelines using Apache Beam (Dataflow), handling customer data ingestion, automated backfill, and late-data recovery, while establishing robust SLAs and SLOs.
- Collaborated closely with product and business teams to define a key cost-per-metric KPI used directly in customer contracts; led engineering teams in achieving cost reductions from \$6.00 to \$1.68 per metric per month, establishing corresponding KR monitoring practices.
- Designed and built machine learning infrastructure enabling real-time scoring of gradient-boosted regression models, significantly improving customers' offline quality outcomes.
- Developed the initial MCP backend powering Oden AI's chat product, including sophisticated entity resolution and automated query composition to answer complex manufacturing queries.
- Published and presented peer-reviewed work on ML infrastructure at ICDM 2019; co-authored a Google Cloud blog post on Dataflow and BigQuery; and regularly spoke at Apache Beam Summits (2019–2023).

Columbia University

New York, NY

INSTRUCTOR, PRE-COLLEGE PROGRAM

Oct 2024 – Present

- Designed and taught both levels of Columbia's Data Science and Machine Learning course for high school students preparing to pursue CS or DS at the university level.
- Developed a full curriculum covering Python fundamentals, statistics, data analysis, NLP, classical ML with scikit-learn, neural networks with PyTorch, and introductory LLM-based applications.
- Authored over 470 slides and 34 hands-on assignments to support project-based learning in classes of 15–25 students.

Chartbeat

New York, NY

SENIOR SOFTWARE ENGINEER AND TEAM LEAD

June 2013 – July 2017

- Architected and led development of the backend for Chartbeat's "Heads Up Display," a real-time analytics tool processing over 100M click events per minute with sub-15s latency. Work included real-time data imputation, batch-mode historical benchmarking, and estimating clickthrough rates from sparse event data.
- Served as team lead and Scrum Master, guiding a 9-person engineering team to successfully launch and maintain core analytics products.
- Built Chartbeat's initial Kafka consumers in Clojure, maintained Chartbeat's open source Clojure-Kafka libraries, and established internal best practices for using Clojure and Kafka at scale.

Languages Python, Java, Clojure, Go, SQL

Tools Apache Beam, Airflow, PubSub, BigQuery, Kubernetes, PostgreSQL, MongoDB, Redis, BigQuery, AWS, GCP

Education

Rutgers University – School of Arts and Sciences

New Brunswick, NJ

B.S. IN COMPUTER SCIENCE

2013













- President, Undergraduate Student Alliance of Computer Scientists
- Lead Organizer, HackRU Hackathons (2012–2013)
- Recitation Assistant, Mathematics Department; Lab Assistant, Computer Science Department
- HackNY Fellow (Spring 2012); Software Engineer Intern at Branch.com

Talks and Publication


ACADEMIC PUBLICATION

- 2019 **ICDM, Mimir: Building and Deploying an ML Framework for Industrial IoT** Beijing, China
A framework for building and deploying Industrial IoT ML models and the outcomes of a predictive model in Plastics Extrusion.  



CONFERENCES

- 2025 **Apache Beam Summit, Real-Time Predictive Modeling w/ Beam, MLServer, and MLFlow** New York, New York
Real-time scoring and ML infrastructure of predictive offline quality test models. 
- 2023 **Apache Beam Summit, A Beginners Guide to Avro and Beam Schemas** New York, New York
Methods for using Avro with PubSub, BigQuery, and Dataflow for both type safety and cost reduction.  
- 2022 **Apache Beam Summit, Detecting Change Points in Apache Beam** Austin, Texas
Methods for detecting change-points in a streaming system considering the impact of lateness, order, and sparsity.  
- 2021 **Apache Beam Summit, Batch-Mode Late-Data Processing for Streaming Pipelines** Online
Using Beam's combined batch and streaming model to build a lambda architecture for processing late data.  
- 2017 **PyCon Japan, How (and Why) We Speak in Unicode** Tokyo, Japan
A history of encoding across the world, the impact of Unicode, and why it's so hard in Python.  
- 2016 **Clojure/West, One Million Clicks per Minute with Kafka and Clojure** Seattle, Washington
How to use Clojure to write Kafka consumers and producers and lessons learned from building real-time analytics applications.  
- Meetups, NYC Apache Airflow Meetup, NYC Timeseries Meetup, NYC IoT Meetup, NYC Apache Beam Meetup** New York, New York
In addition to conferences, I regularly speak at meetups. Slides for most of my talks can be found at github.com/x/slides. 

BLOG POSTS

- 2023 **Google Cloud Blog, Built with BigQuery: How Oden provides actionable recommendations with network resiliency to optimize manufacturing processes** Online
A blog post Co-Authored with Google Cloud Director of Applied AI Engineering on how Oden uses BigQuery, PubSub, and Dataflow to build resilient real-time machine learning pipelines. 

COURSES

- 2025 **BIGD204, Data Science and Machine Learning 2** Columbia University
Intensive overview for students with programming experience of Data Science, Machine Learning, and an introduction to LLMs, emphasizing real-world data and building applications with traditional ML models and pretrained LLMs (20 hours). 
- 2024 **BIGD103, Data Science and Machine Learning 1** Columbia University
Intensive introduction to Python, Data Science, and Machine Learning for students without prior programming experience (60 hours). 

Hobbies and Interests

Outside of work I build and maintain side projects such as [a party planning app](#), [a Pokémon TCG Pocket draw simulator](#), and [an automated "12 Seasons of NYC" meme site](#), which was [referenced in official MTA service changes](#).

When I'm not writing code I write and run a multi-year [Dungeons & Dragons campaign](#) as a Dungeon Master and am an avid hiker and long-distance cyclist, once biking [from NYC to Montreal](#).