

Devon Peticolas

SOFTWARE ENGINEER

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

About

I specialize in architecting real-time and historical data systems, productionizing Data Science and Machine Learning algorithms, and building data-driven products. I believe in leading by teaching and have presented at meetups and conferences about Data Engineering and Data Science in NYC and internationally.

I am seeking a lead individual contributor engineering position focused on challenging problems that deliver direct user value.

Work Experience

Oden Technologies

New York, New York

PRINCIPAL SOFTWARE ENGINEER, DATA SCIENCE TEAM

Oct 2017 - Present (5 years)

- Built multiple systems to productize Data Scientist's research into customer-facing production-ready features, including our reports system, which builds customer-specific daily-updating reports from Jupyter notebooks written by Data Scientists.
- Led a team of Data and ML Engineers to build a real-time customer-data ingestion and enrichment system on top of Apache Beam with a lambda-architecture-style automatic batch recovery system to account for late data and outages.
- Led a team to productionization random-forest and regression models into historical and real-time recommendation systems.
- Published and presented a paper at ICDM 2019 on our ML infrastructure and models and their performance with customer data.
- Reduced our marginal cost-per-customer by 35% over one quarter by modeling and reducing our cost per data-point ingested through a series of strategic changes to our streaming architecture.

Chartbeat

New York, New York

SR. SOFTWARE ENGINEER AND TEAM LEAD

June 2013 - July 2017 (4 years)

- Architected and spearheaded the team that built our real-time backend for our "Heads Up Display," a real-time homepage analysis tool that tracked 100M click events per minute and delivered live performance analysis in less than 15s. This involved real-time imputation of gaps in data, batch modeling for historical benchmarks, and real-time estimation of clickthrough rate from sparse events.
- Became a certified Scrum Master and adopted a Scrum process to lead a 9-person team to a successful product launch after a year of setbacks. I applied the same approach to my next team to rebuild our reporting product.
- Was the primary maintainer of our Clojure-Kafka libraries and led our Clojure and Kafka best practices.

Languages Python, GoLang, Java, Clojure

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

Education

Rutgers University - School of Arts and Sciences

New Brunswick, NJ

B.S. IN COMPUTER SCIENCE AND MATH

Class of 2013

- President of The Undergraduate Student Alliance of Computer Scientists.
- Chief Organizer of the Spring-2012, Fall-2012, and Spring-2013 HackRU Hackathons.
- Recitation assistant for Math department and lab assistant for the Computer Science department.
- HackNY Fellow of Spring 2012; Software Engineer intern for Branch.com working in Ruby on Rails.

Talks and Publication

ACADEMIC PAPERS

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

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

In addition to conferences, I've spoken at the **NYC Apache Airflow Meetup**, the **NYC Timeseries Meetup**, the **NYC IoT Meetup**, and the **NYC Apache Beam Meetup**. Slides for most of my talks can be found at github.com/x/slides.