

Frank Cleary

Technical leader focused on scalable systems for analytics, search, and machine learning

frank@frankcleary.com • frankcleary.com • github.com/frankcleary

Highlights

Java: Microservices, implementing query engines and languages, machine learning

Big Data: Elasticsearch, Kafka, Spark, Hadoop, AWS (EC2 and related services)

Python: Pandas (minor contributor), Numpy, Matplotlib, nltk, Gensim

Other: ANTLR, MongoDB, SQL, Linux, Git, gradle, security (personal sites use https), \LaTeX

Data Science Bytes: Website for data focused tutorials and tips. 4k monthly unique visits.

- [Spark Streaming Tutorial](#). Instructions to get started with streaming big data
 - [Recommender for Pelican](#). Presented as a lightning talk at the SF Python Meetup in February 2015
 - [D3 + SQL on AWS](#). Dynamic graphs powered by flask, D3, and SQL, hosted on AWS
-

Education

Ph.D., Biophysics

NSF Fellow. frankcleary.com/research

University of California, Berkeley

2008 – 2014

B.A., Physics

Cum laude, with distinction in all subjects

Cornell University

2002 – 2006

Experience

AppDynamics (acquired by Cisco)

Staff Software Engineer

Senior Software Engineer

Software Engineer II

San Francisco, CA

2/2018 – Present

9/2017 – 2/2018

1/2016 – 9/2017

- Technical lead for ADQL, the query language and engine that underpins AppDynamics' analytics product
- Responsible for technical direction and project planning for the ADQL team
- Transformed ADQL language modules, enabling query optimizations (up to 50x improvement) and new language constructs (such as the HAVING clause)
- Designed and led implementation of a query engine to support operations on streaming data
- Created an async RPC framework to allow service method level request prioritization and load shedding within AppDynamics' existing Java server framework
- Developed support for math operators, aggregation filters, and sorting in ADQL. These features were demoed a week after their release at the CEO's keynote at AppDynamics' annual user conference, two months before our planned IPO and eventual \$3.7B acquisition by Cisco
- Organized a hackathon team around new search data structures, won the infrastructure category (50 teams total)

Index (acquired by Stripe)

Member Technical Staff

San Francisco, CA

6/2014 – 11/2015

- Implemented production recommender systems and conversion rate tracking leveraging Spark and MongoDB
- Developed monitoring and alerting of the company's services, greatly increasing visibility into production issues
- Delivered data analysis projects using Hadoop and the Python data science stack

Seven Networks

Data Scientist Intern

San Carlos, CA

6/2013 – 8/2013

- Developed data analysis software in Python, including a service to display key internal metrics
- Analyzed A/B test data to provide summaries and recommendations on test design

New York City Public School System, PS/MS 95

Middle School Science Teacher

Bronx, NY

6/2006 – 6/2008

- Selected and trained as part of the Teach for America program
- Taught classes of 30 middle school students at PS/MS 95 in Bronx, NY