Eric Schles

[ericschles@gmail.com](mailto:ericschles@gmail.com)

Languages: Python, Javascript, Java, R, HTML/CSS

Web Technologies: Flask, Django, Express, Node, Pelican

Databases/Backend: Postgres, Oracle, MySQL, SQLite, MongoDB

Cloud technologies: Cloud.gov, CloudFoundry, AWS

**Experience**

Data Scientist

Innovation Specialist - 18F August 2016 – Present

*Projects:*

* CALC: I work as the data scientist for CALC. In this project I work on a series of timeseries visualizations, prediction models, and innovations. I’m also working on, you saved x with calc, an internal white paper about whether calc is saving the government money.
* analytics.usa.gov: Deployed this from AWS East/West to GovCloud and cloud.gov
* Open Opportunities : Worked on backend issues and bug fixes for this project. First exposure to cloud.gov
* Discovery: Deployed from AWS East/West to GovCloud.
* SkillShare: An internal 18F tool for sharing current skills 18Fers might want to share or learn. Worked on Developing and deploying this flask app to govcloud (still being developed)
* Python Intro class: Teaching an intro to python class to other non-technical 18Fers, covers data structures, functions, classes, variables, debugging, basic web dev, and making python fast via async interfaces.
* Python Data Science Class: I’m working on an open source book on machine learning and I’m dog fooding the course material with 18Fers, covers topics from discrete to continuous models. Some highlights: PGMs, Linear Regression, Gaussian Regression, Decision Trees, SVMs, Neural Networks, Kaggle problems and strategies.
* Data Services Working Group: I help lead the data services working group where we discuss best practice and common strategies for working on data projects in government.
* Selling New Contracts: I’ve developed a few leads for 18F. Nothing has lead to any contracts as of yet, but I’ve started some conversations.

Adjunct Professor

Python Lead Instructor

NYU – January 2017 – Present

I teach a year long Python intensive at NYU. In this class I take students from nothing to web development. Covering data structures, algorithms, classes, functions and more. Web development is done using Flask.

Data Scientist

Digital Services Expert – United States Digital Service, VA Team October 2015 – Present

*Projects:*

* Vets.gov Search: I manage vets.gov search. For this project I wrote a custom indexer which connects to i14y’s search.
* Caseflow Reports: This tool incorporates data visualization, bootstrap, and other design elements to create a simple visual experience for the caseflow system. Caseflow is used for the appeals process at the VA.
* Caseflow Analysis : This project also works with caseflow and performs analysis on expected wait times as well as determines what parts of the appeals process are taking the longest. This project takes into account data visualization, graph based algorithms, and descriptive statistics.
* Database version control system : This project was used to take snapshots of databases over time, since data was regularly destroyed and updated. By versioning the database, time series analysis and further reporting could be conducted.
* Automatic merging : This project involved a number of complex database operations, specifically sophisticated joins, data cleaning, and normalizations. It was used to combine 17 different databases into one.

Data Scientist - Demand Abolition July 2015 – October 2015

In this role I worked on a variety of projects and consulting in automation, prediction, analysis of data, and drawing clear lines between national trends in human trafficking.

Senior Analyst – Manhattan District Attorney October 2014 – August 2015

Human trafficking Response Unit

*Projects:*

* Anomaly Detection: During this project I created a tool that finds and ranks financial records as possible instances of money laundering.
* Search Query Optimization: This tool makes use of text classification and the textrank algorithm to determine optimal query terms to search for trafficking within police reports.
* Automated ETL : During this project I developed a general purpose tool to parse unstructured data, transforming the relevant data into a machine readable document. Using this tool, analysts can transform PDFs, HTML documents, and text documents into Excel documents as well as prepare them for ingestion into a proprietary data-store.
* Facial Recognition platform : For this project I developed a system that can access all the pictures on a file system as well as from a database and compare them against a newly ingested phone.
* Tip/Lead generation : A tool that takes in past records of human trafficking and then based on key factors generates new leads from data scraped from the internet. Natural language processing, specifically sentiment analysis, document similarity analysis, facial recognition and keyword/frequency analysis were used.
* Investa\_gator: Scrapes backpage, to automate investigations. The system does entity recognition to find phone numbers, names, email addresses, and other hard attributes.
* HoneyPot operations: Created a number of tools to assist investigators in capturing suspected individuals.
* Speech To Text translation: Cleaned and manipulated a number of audio recordings and created automatic transcriptions.

* Email Parser: Aided in the creation of an email parser that would parse hundreds of thousands of emails for specific keywords. This parsing was done in conjunction with Hadoop.
* Website forensics automation: this tool creates a local version of a website in order to have a complete picture of what the website said over time.
* Picture Toolkit: A fully automated picture analysis website and framework. Developed in house, custom for the uses of the Human trafficking response unit. Capabilities include, stenographic encryption in pictures, exif manipulation and removal, and watermarking.

Contract Researcher – New York University, New York, NY January 2014 – Present

*Projects:*

* Developed Research Website : During this project I developed the website used to recruit participants for the study. The website was developed as a custom solution using python, javascript, html, and css.
* API integration : At present I am developing software to gather data from facebook, twitter, whatsapp, sms text messages, google+, and a few other social media sites. This data will be used to analyze the textual data from participants in the study.
* Machine Learning Consulting: Aided research professors write code in python. Specifically with integrating a custom linear regression algorithm that made use of scipy.optimize, numpy arrays and scipy functions.

Adjunct Professor – New York University, New York, NY December 2013 – Present

* Courses taught: Applied Data Science, Introduction to Data Structures, Algorithmic Problem Solving.

Developer Evangelist – Syncano, New York, NY November 2014 – June 2015

* Developed and improved api documentation for python and javascript api's
* Presented weekly on new and trending technologies as well as incorporated these technologies into the Syncano platform
* Wrote and developed several technical blog posts, with original code and writing. Blog posts reside at: http://www.syncano.com/blog/
* Built a successful community of full stack developers with over 500 members.

Data Scientist – Lyte, New York, NY March 2014 – September 2014

*Projects:*

* Automated Pricing : The goal of this project was to construct a model that determines a fair price for a ticket on the second hand market.
  + Accomplishments:
    - Wrote a number of fully automated web scrapers to pull sports data to determine team popularity
    - Designed and implemented the model, which updates dynamically based on information from the webscrapers and proprietary data sources.
    - Algorithms used: K-means clustering, Support Vector Machine
* Analytics Dashboard : The goal of this project was to develop analytics to improved user experience and allowed for great decision making power from a strategic perspective.
  + Accomplishments:
    - Implemented a dashboard that logs all actions on the website
    - Shows user interaction in a tabular form

Consultant – Skim.me New York, NY April 2012 – September 2012

*Projects*:

* Incorporated software engineering best practices into the product, which improved software scalability and allowed for a faster time to market.

Data Miner – Brooklyn Boulders, Brooklyn, NY July 2011 – Jan 2012

* Discovered salient facts about the online community to aid in future marketing campaigns.

**Pro Bono Work**

The News Literacy Project

In this project I lead a team of 45 – 100 developers in one off “workathons” once a month. Developers come and work on stuff day of. So far we’ve moved their infrastructure to github pages. And are currently working on a system to classify news as fake or likely real.

The Manhattan DAs HTRU

After leaving the human trafficking response unit I’ve continued to collaborate with them. Giving them prototypes and free labor.

Contra Costa DAs Human Trafficking Team

Much like NY, I help out in contra costa from time to time as well. Typically I’ll build prototypes, small projects and give advice.

**Education**

Masters of Science in Computer Science – New York University Sept 2011 – Dec 2013

Honors:

National Honors Scholarship Fellow: Scholarship for Service

Dean's list three semesters

Two Merit Based Scholarships for high grades

Masters of Science in International Economics with a focus in Econometrics Sept 2009 – May 2011

Honors:

Dean's list three times

Projects:

Designed and Developed an algorithm to determine whether businesses pollute within the bounds of the law

Bachelors of Science in Mathematics and Economics – SUNY Buffalo Sept 2005- May 2009

Honors:

Cum Laude

Dean's list eight semesters

Projects:

Developed a novel financial derivative to price stocks