

# ZoëFarmer

Data Scientist

## About

345 South 40th Street  
Boulder, Colorado,  
80305

(719) 233 3514  
dataleek.io

twitter/TheDataLeek  
zoe@dataleek.io  
github/thedataleek

## Programming

### Expert

Python, LaTeX

### Intermediate

SQL, R, Go, Bash,  
Regex, HTML, CSS

### Novice

Scala, Javascript

## Tools

### Mathematics

Probability/Statistics

### Frameworks

Spark, Numpy, Scipy,  
Flask, CI Tooling

### Visualization

D3, Graphviz,  
NetworkX, Matplotlib

### Infrastructure

Nginx, Apache,  
Docker, LAMP

## Education

B.S. in Applied Math  
CompSci Minor  
C.U. Boulder

## Experience

- 03/16 - Now **Nidhogg Consulting, Boulder, Colorado** Data Scientist / Software Developer  
*Co-founded Nidhogg Consulting, a consulting security firm as Chief Data Scientist and Software Engineer. Responsibilities include project management, full stack web development, data science projects, and job acquisition. Technologies used include Python, Flask, jQuery, Bootstrap, Docker, CI Tooling, and Javascript.*
- 09/15 - 05/17 **University of Colorado, Boulder** Dispersive Hydrodynamics Research Assistant  
*Implemented data quality metrics using machine learning algorithms. Also responsible for much of the lab's data analysis projects. Developed and implemented a massively parallel simulation framework to analyze behavior of soliton gasses in one-dimensional systems. Presented at SIAM 2016 and SIAM 2017. Technologies used include Python, Numpy, Scipy, Scikit-Learn, Matplotlib, Multiprocessing, Matlab, SQLite, and Supercomputer Design.*
- 05/16 - 08/16 **Khan Academy, Mountain View, California** Data Science Intern  
*Developed new features and maintained khanacademy.org as well as enhanced the built-in site metrics. Assisted in the SAT research project with Collegeboard using investigative analytics to accurately describe the dataset. Also developed a generalized learning model to assist with student growth and projection estimates for the SAT practice tool. Technologies used include Python, Flask, React, PyData Stack, and Data Visualization tools.*
- 06/15 - 08/15 **Shutterstock, Denver, Colorado** Data Scientist Intern  
*Created a Apache Spark application to handle large scale ETL tasks on top of their Hadoop platform, as well as deployed several big data solutions.*
- 04/14 - 09/14 **SpotXchange, Westminster, Colorado** Data Scientist Intern  
*Developed an automatic Tableau Server reporting tool based on the semi-public API. Analyzed time-series data in an effort to identify similarities. Created a probabilistic model to analyze single-threaded schedule optimizations.*
- 05/13 - 08/13 **MITRE, Colorado Springs, Colorado** Data Visualization Intern  
*Created network data visualizations based on packet transfer data using common visualization technologies. Implemented encryption section for 2013 CTF.*
- 08/12 - 05/13 **Snow and Ice Data Center, Boulder, Colorado** Data Management Assistant  
*Large scale ETL involving terabytes of research data. Involved application of optimized Python and bash scripting.*

## Projects

**Boulder Python Meetup Coordinator**

[meetup.com/BoulderPython](https://meetup.com/BoulderPython)

**Latex Forum Moderator**

[reddit.com/r/LaTeX](https://reddit.com/r/LaTeX)

**Enforce**

[github.com/RussBaz/enforce](https://github.com/RussBaz/enforce)

**Black Arrow**

[github.com/thedataleek/black-arrow](https://github.com/thedataleek/black-arrow)

**Twitter Panic**

[github.com/thedataleek/TwitterPanic](https://github.com/thedataleek/TwitterPanic)

**Kobayashi**

[gitlab.com/thedataleek/koabayashi](https://gitlab.com/thedataleek/koabayashi)

**SnekCount**

[gitlab.com/thedataleek/snekcount](https://gitlab.com/thedataleek/snekcount)

**Kobayashi**

[gitlab.com/thedataleek/kobayashi](https://gitlab.com/thedataleek/kobayashi)