

# Joseph Haaga

## Developer & Data Engineer

2507 17th St NW Apt B, 20009, Washington, D.C. US

<http://joehaaga.xyz> [haaga.joe@gmail.com](mailto:haaga.joe@gmail.com)

(202) 525-6891

[joehaaga](#) [joehaaga](#)



### SKILLS

---

#### Web Development

html | css | lamp stack | mean stack |  
reactjs | nodejs | jquery | pimcore |  
drupal 7/8 | wordpress | 508 compliance

#### Ops/Infra

docker | terraform | ansible | kubernetes |  
aws

#### Software Development

tdd | oop | functional programming |  
design patterns | algorithmic analysis |  
postgres | mongodb | neo4j

#### Data Science

pandas | tensorflow | numpy | nlp |  
classification | regression/linear modeling |  
clustering/unsupervised techniques |  
sql/nosql | elasticsearch & logstash |  
tableau | arcgis

### WORK EXPERIENCE

---

#### Senior Data Engineer at Capital One August 2019 - Current

<https://capitalone.com>

Capital One Financial Corporation is an American bank holding company specializing in credit cards, auto loans, banking, and savings accounts, headquartered in McLean, Virginia with operations primarily in the United States

- Supported efforts to achieve compliance with enterprise Cloud security mandates
- Automated reporting for cloud infrastructure vulnerability scans
- Led assessment & remediation of difficult cloud modernization efforts (Terraform, Ansible, boto3)

#### Data Scientist at Tylor Data Services October 2018 - August 2019

<http://tylordata.com/>

TDS continues to deliver data science consulting, big data analytics, and custom computer programming services that are proven to address business challenges for government agencies, non-profit organizations and private corporations.

- Led development of a graph analytics tool (React, Flask, Neo4j, Docker) for fraud pattern

detection in financial data graphs.

- Supporting research in applying novel deep learning techniques (RNN embeddings, Tensorflow, graphwave) to elicit graph embeddings.

## Data Scientist at Redhorse Corporation June 2018 - October 2018

<https://redhorsecorp.com>

Redhorse Corporation, a program management and technology services company, provides integrated solutions in environmental consulting, energy and climate change, information technology, and intelligence.

- Designed and implemented a image recognition application for environmental science use case.
- Experimented with different convolutional neural network architectures to maximize F1 score.
- Built custom sklearn Pipeline components to take advantage of OpenCV preprocessing techniques.
- Developed web-based utility resulting in a 10x decrease in time spent labelling training data.

## Full Stack Developer at Grafik August 2016 - June 2018

<https://grafik.agency>

Grafik is a branding agency focused on crafting digital brands for B2B, B2C, Nonprofit and government clients worldwide.

- Led architecture and development of CMS-driven solutions for clients like Strayer & Global Automakers, using Drupal 8 and Pimcore 5.
- Contributed to the architecture and development of marketing websites for brands like Strayer, Ligado, Politico, Merritt Clubs, JBG and EYA. Developed on Drupal 7, Pimcore 4/5, and Wordpress.
- Drastically reduced development timelines by building internal tools to streamline internal processes (structured QA spreadsheet with Slackbot integration to close developer-designer feedback loop and closely track live performance/completeness metrics). Built using Google Apps Script, Google Sheets, and Slack Webhooks.
- Delivered customer persona insights by applying data mining & visualization techniques to client CRM & transactional data. Performed analysis using Tableau and Anaconda toolkit (Seaborn, Pandas, NumPy).

## Lead Architect at GWU Data Science Institute November 2017 - May 2018

<https://gwu.edu>

Selected to be a founding member of GWU's DSI, aiding in bootstrapping the institute as well

as supporting technical initiatives.

- Led deployment efforts in standing up a cloud-hosted information sharing environment for Virginia state agencies, alongside ODNI and AWS
- Mentored students in Data Warehousing classes, providing guidance on their semester projects

## Developer at Gray Street Solutions May 2015 - July 2016

<https://graystreetsolutions.com>

Gray Street Solutions is a boutique, digital-first creative agency, founded on the principle of bringing a “Fearlessly Agile” approach to online communications consulting, open-source web development, and social media campaign management.

- Led development of a cross-platform mobile application to track a brand’s social media campaigns, using Apache Cordova and Adobe PhoneGap.
- Developed and maintained digital presences for a variety of trade organizations, defense companies, and nonprofit clients. Developed on Wordpress and maintained on WPEngine managed hosting.

## Jr. Solutioneer at Eastport Analytics May 2014 - August 2014

<https://www.eastportanalytics.com/>

Eastport Analytics is an analytics consulting group applying advanced analytics to uncover valuable insights and drive measurably better mission outcomes.

- Developed automated spreadsheet parsers in Java 7.
- Merged GIS shapefiles with political ad spend data to determine campaign efficacy in Designated Market Areas.

## Developer at @HomeRehab May 2013 - April 2014

<https://gust.com/companies/homerehab>

@HomeRehab designs exercise devices and software to improve the continuum of care related to falls among the aging population.

- Prototyped IoT orthopedic rehabilitative device on Arduino hardware in a dialect of C.
- Developed games to be played with the device to aid in gamification of proprioceptive rehabilitation exercises.
- Implement wireless modules (XBee), LCD screens, & various sensors/actuators within I/O constraints of a development board.

## Coordinator & Volunteer at Grafik Gives

August 2017 - Current

<https://grafik.agency>

Introduce potential initiatives and coordinate CSR events.

- Raised >\$1000 for AIGA, providing scholarships for underrepresented design students in the DC area.
- Developed spreadsheet solution for long-term initiative planning, enabling tracking of the organization's impact on local, regional, and national levels.

## Kitchen Volunteer at So Others Might Eat

April 2011 - Current

<http://some.org/>

Serving meals to marginalized people in the DMV area.

### EDUCATION

---

## M.Sc. Computer Science at Georgia Tech

2019 - Current

Major courses: [software analysis and testing](#) | [software architecture and design](#)

## M.Sc. Data Science at The George Washington University

2017 - 2018

Major courses: [data science](#) | [data warehousing](#) | [data mining](#) | [visualization of complex data](#) | [machine learning i](#) | [machine learning ii](#) | [time series analysis](#) | [cloud computing](#) | [bayesian methods](#) | [member of graduate consulting club](#) | [member of mba business analytics club](#) | [president of ms data science student council](#)

Grade: 3.92

## B.Sc. Computer Science at The George Washington University

2012 - 2017

Major courses: [circuit theory](#) | [bioinformatics](#) | [algorithms](#) | [software engineering](#) | [database systems](#) | [systems architecture](#) | [operating systems](#) | [machine learning](#) | [systems programming](#) | [artificial intelligence algorithms](#) | [statistical nlp](#) | [deepquest ibm watson \(essentially a supplement to stat. nlp\)](#)

### LANGUAGES

---

Python

BASH

advanced

Javascript  
intermediate

Java  
beginner

intermediate

PHP  
beginner