Hugh Wimberly

hugh.wimberly@gmail.com
(505) 750–4844

2005 Rose Street
Berkeley, CA 94709

Relevant Experience

Headspsace Health 11/2020–6/2022

Senior Data Engineer

- Supported the infrastructure and data needs for a team of 6 ML engineers and 2 data analysts
- Built maintainable machine learning pipelines orchestrated with AWS CDK
- Designed HIPPA-compliant de-identification schemes to satisfy cross-functional analytics needs
- Created an efficient and secure n-gram aggregator for detailed trend analysis

Project Deepform *(nonprofit supporting data extraction for journalism)* 6/2020–11/2022

MLOps Engineer

- Migrate project to using Docker and docker-compose for reproducible builds
- Automated code-quality checks, integration tests, and image deployment with GitHub Actions
- Reduce single-instance traning and inference times by 90% by using Parquet and Modin
- Improve F1 score by 7% using multitoken labeling and longer prediction sequences

Toptal (Data Engineering Consultant) 5/2018–5/2020

fortune 500 insurance company

- Advised on transition from on-prem to cloud (AWS) hosted data and services
- Simplified and streamlined transition plan, reducing third-party dependencies
- Designed kafka-based interop between extant mainframes and new elastic ECS services

fortune 100 healthcare company

- Developed methodology for comparing performance of initiatives run by different teams
- Designed data architecture for major cross-team reporting dashboard
- Guided data science team in adopting best-practices

major canadian retail group

- Moved core analytics database from SQL Server to Azure Data Warehouse
- Optimized bottleneck operations, improving analytics latency from 20 hours to 40 minutes
- Guided team adoption of Docker, CircleCI; led testing and automation initiative

fortune 500 restaurant group

- Migrated data science workflow to Azure Data Warehouse (previously self-managed)
- Reduced analytics pipeline latency from 30 hours to 6 hours and improved reliability
- Advised and assisted in transition from pure Python to a PySpark workflow

education startup

- Reduced the AWS bill by 80% by applying DynamoDB best practices
- Improved the page-load speed 5x by prefetching and prewarming serverless APIs
- Eliminated the need for additional Redis hosting and caching
- Reduced the API latency by 70% by increasing parallelism

Udacity 2/2015–8/2015

Senior Backend Engineer

- Optimized backend to reduce CPU usage by 70% and decrease AppEngine bill by $30K/mo.
- Migrated legacy monolithic AppEngine site to AWS microservices
- Internationalized and localized main website for India launch
- Replaced payment backend to add Stripe support

Google, *Knowledge Graph & Machine Intelligence* 2/2012–2/2015

Software Engineer III

- R&D work on Google-scale in-memory distributed graph database
- Wrote two optimizing query compilers (for two source languages)
- Created visual debugger for inspecting graph diffs
- Led project test infrastructure improvements (recognized with Spot award)

New Mexico Institute of Mining and Technology 8/2009–12/2011

Instructor, CSE 221 Computer Systems Organization

- Set curriculum and syllabus to meet new course objectives
- Created 4 team projects, 6 solo projects, 2 midterms and a final
- Taught a lecture and lab for 33 students
- Responsible for all grading—wrote import scrips and autograders to ease project grading

Los Alamos National Laboratory, *Advanced Computing Solutions (ACS)* 01/2011–08/2011

Scientist II

- Designed and prototyped a 100,000–node fully decentralized Command & Control framework
- Wrote firmware patches to update network packet capture software for new 10G interfaces
- Carried out experiments to compare architectures for a custom petabyte scale database
- Assisted in the design and analysis of a 10,000–subject phishing experiment

Idaho National Laboratory, *Nuclear non–proliferation team* 6/2010–8/2010

Computational specialist

- Wrote high-performance inverse spectroscopy GPU code (CUDA) for a novel radioisotope detector
- Worked with a team to develop statistical analyses of fuel diversion scenarios

New Mexico Computing Applications Center 6/2008–12/2008

Simulation team lead

- Designed an MPI-parallel traffic simulator for the New Mexico Department of Transportation
- Taught a summer high school programming course

Boost C++, *Google Summer of Code* 5/2007–8/2007

Open-source developer

- Extended the Boost.Regex C++ library to allow recursive matching and named sub-expressions

Microsoft Corporation, *Anti-phishing team* 5/2006–8/2006

Sofware Development Engineer Intern

- Constructed prototype server frameworks to evaluate replacements for IE's anti-phishing service
- Discovered and repaired a latent security vulnerability in the Microsoft anti-phishing servers

Geophysical Solutions 5/2003–8/2003

Sofware engineer

- Wrote code to render 3-D maps of underground magnetic fields
- Worked with electromagnetic sensors in the field to gather test data
- Rewrote and updated FORTRAN 77 geodetic software in C++

EDUCATION

New Mexico Institute of Mining and Technology (NMT)

pursued Ph.D in Computer Science, no degree awarded 8/2007–12/2011

- Concentration: Information Security and Assurance

Rensselaer Polytechnic Institute (RPI)

B.S. in Computer Science 8/2003–5/2007

- Concentration: Bioinformatics