CONTACT

London, UK

07568 199812

charles.d.waller@gmail.com

GitHub

LinkedIn

View My PwC Career Article

SKILLS - SPECIALISMS

Rest APIs

Microservices

Distributed Systems

Web Scraping

Natural Language Processing

Predictive Modelling

Data Engineering

Behaviour Driven Development

Test Driven Development

SKILLS - BACK-END

Golang

Python

Bash

MySQL

PostgreSQL

BigQuery

Elasticsearch

NoSQL

Redis

Web Scraping

PySpark

Charlie Waller

I am a Software Engineering Manager within PwC's Artificial Intelligence team, passionate about technology and personal development.

From starting as a Consulting Apprentice at PwC 6+ years ago, I've proactively pursued my passion which has seen me specialise as a Software Engineering Manager in PwC's Artificial Intelligence team. I am now looking for a new challenge within a technology business to expand my skill set and grow professionally.

Outside of work, I'm regularly pursuing projects / learning to expand my skills and challenge myself further. Away from the desk, I love sailing, surfing and music.

PwC PROJECTS

Establishing AI Team Engineering Team

Sep 2020 - Present

Context: In addition to project work, over the last 12 months I have been working to establish a central engineering capability within the AI team. This is with the intention of fostering best practice, code reusability and engineering culture within the wider team.

Role: Mid-Level Software Engineer

- Taking an active role in recruitment, including CV screening and facilitating technical interviews
- Providing training and development to junior engineers through paired programming and "Show & Tell" sessions
- Mentoring junior engineers with regards to PwC career progression
- Supporting the development of a central technology career path within PwC

Outcomes: Since September 2020, the AI Engineering team has grown from 4 to 15 people, supporting 4 projects with a high level of code reuse - increasing code quality and reducing delivery time on engagements.

Recognised for my contribution towards establishing the Engineering Team, I was promoted to Manager in January 2022.

Document NLP Platform

Sep 2020 - Present

Context: Built a Document NLP Processing platform applying Machine Learning and Advanced Analytics techniques to large corpuses of data. This is used by project teams to save time on repeatable analysis.

Written in **Golang**, this application is built using a microservices architecture, hosted on **Kubernetes** using Google Kubernetes Engine (**GKE**).

Role: Mid-Level Software Engineer

- Writing resilient, extensible and fully tested REST API services in **Golang**. For example, implementing a text extraction service using external OCR APIs
- Triaging bugs, diagnosing the root cause and implementing tested

Gherkin

Prometheus

Grafana

GCP

SKILLS - DEVOPS

Docker

Kubernetes

Azure CICD

Pipelines

Terraform

SKILLS - FRONT-END

React

JavaScript

HTML

CSS

SKILLS - SOFT

Agile / SCRUM

Training delivery

Meeting facilitation

Collaboration

QUALIFICATIONS

Python for Data Science and Machine Learning, Udemy - May 2019

Certificate in Management Consulting and Diploma in Business Skills

A-Levels - Business, Economics, Product Design (A*, A, A)

fixes, enabled through feature flags.

- Mentoring more junior engineers in the team in the form of paired programming, detailed code reviews, and sharing of good practice
- Making design decisions around implementation of core components, considering impacts on scalability, resilience and cost
- Deploying all application services to the **Kubernetes** environments, building CICD pipelines, managing container registries (**Docker**) and enforcing Policy Based Deployment.

Outcomes: Using **Prometheus** and **Grafana** for monitoring we can see that the platform is used by multiple client teams across PwC, and has saved hundreds of man hours of repeatable analysis.

Code Generation SQL REST API *Jul 2021 - Sep 2021*

Context: Built a REST API that uses historic PwC project data to price new engagements more effectively. This was to combat inconsistent pricing across engagements.

My approach to this project was to generate the API code automatically based on a specification file, allowing for quick iteration when business requirements changed.

Role: Mid-Level Software Engineer

- Writing the code generation CLI using **Golang** templates, generating fully tested API code, and working with the business team to iterate on the specification
- Building the data mart for the Deals use case in **PostgreSQL**. This included joining data from a number of internal data sources, performing feature engineering to calculate desired metrics for filtering, and applying relevant indexes to optimise API queries
- Working with the line of service to deploy the application, implementing the required authentication and packaging the application and its dependencies using **Docker**

Outcomes: The API is used by the Deals line of service to inform project pricing. Due to the code generation approach, the same codebase has been used to build APIs for 3 other projects

Climate Change Text Classification Model *May 2020 - Sep 2020*

Context: Built a text classification model to categorise annual reports based on climate change related content. This is to enable the TCFD to create industry benchmarks for climate change disclosures.

Role: Data Scientist / Data Engineer

- Exploratory data analysis on the raw text data and working with climate change SMEs to refine the model requirements.
- Designing an active learning approach to labelling (using **Prodigy**) which allowed us to increase label volume and utility in the available time
- Built distributed feature engineering pipelines using **PySpark**, **PostgreSQL**, **BigQuery**
- Owned the workstream for model training and iteration, using **scikit-learn**, **XGBoost**, and **SparkML**. Experiment tracking implemented using **MLFlow** for experiment parameter and metrics tracking

- Built the **PySpark** job to distribute the trained models to predict on the corpus of over 300 million paragraphs using **PySpark** and **PyArrow**

Outcomes: The text classification analysis formed the basis of the published 2020 TCFD status report, in which I am referenced.

NHS Clinical Coding Automation *Jan 2019 - Jan 2020*

Context: Built several Natural Language Processing products to improve the efficiency of clinical coding within the NHS.

Role: Junior Software Engineer

- Re-designing the team's primary product to reduce run time from c.1 day to c.10 minutes by improving the efficiency of the **Python** code
- Created an interactive front-end using **Django**, **Bootstrap** and **JavaScript** to accelerate the reviewing of client deliverables

Outcomes: One of the tools that I built was an anomaly detection tool to identify patients with incorrect medical records, based on their medical history and demographic features. For one trust, the impact of using these tools amounted to over £200,000 in additional revenue each month.

ADDITIONAL EXPERIENCE

The Classic Valuer *Feb 2020 - Present*

Context: The Classic Valuer is a valuations tool for the classic car industry. We provide users with accurate, up to date information about the price and price trend of any classic car

Starting as a learning opportunity to expand my skill set, The Classic Valuer has grown into a profitable business with hundreds of weekly users. Studying Business has always been a passion of mine, and The Classic Valuer has enabled me to take an idea from inception into a profitable product

Role: Co-founder & Tech Lead

- Owning the technical decisions around implementation of the site. This includes everything from cloud vendor choice, provisioning of cloud resources, analysis of public APIs and so on
- Development of the key services which underpin the site, written in **Python** and **Golang**
- Maintaining the data ingestion backend, written in **Python** using the **Scrapy** web scraping framework combined with **Selenium** and custom middleware modules.
- Managing DevOps including use of **Docker** for deployment, and **AWS DynamoDB** for a highly available **NoSQL** data store.

Outcomes: After starting out of a passion for Classic Cars and Technology, The Classic Valuer has transitioned from a learning opportunity into a profitable business with c.200 weekly users.