Hal Wall

Data Engineer

ABOUT ME

Analytical, creative, and quietly relentless, I'm a London-based Data Engineer with a background in Theoretical Physics and a deep love for building clean, scalable, and intelligent systems. I work at the intersection of data, machine learning, distributed data architecture, and cloud transformation-helping organisations turn complex information into clarity and action.

I've helped modernise infrastructure for Fortune 500 firms and built data pipelines that power predictive models and business-critical insights.

- Languages: Python, SQL, PySpark
- Data Platforms: Azure Cloud, Databricks, Azure Data Factory, Unity Catalog, Kafka.
- DevOps & Workflow: Git, Docker, CI/CD, Azure DevOps, JIRA, Agile, Scrum.
- Analytics & Visualisation: Power BI, Tableau.

Outside the data world, I'm a devoted reader of fiction and philosophy, a long-distance runner, and an enthusiastic hiker-often found exploring the wilder corners of the UK with my dog. Whether I'm building resilient pipelines or trekking through Glen Coe, I bring the same focused energy: improve continuously, learn deeply, and leave things better than I found them.

CONTACT ME

henry.n.wall@gmail.com

T: 07516 188203

W: wallcache.com

Overview

- **3+ years of Data Engineering experience** in Fortune 500 companies.
- **BSc. Physics with Theoretical Physics** from Imperial College London.
- **Programming proficiency** in Python, SQL and PySpark for data wrangling.

Work

CHUBB INSURANCE, GLOBAL ANALYTICS | *DATA ENGINEER*

NOVEMBER 2024 – CURRENT

Worked at the intersection of data engineering and machine learning within Chubb's EMEA Analytics team and Global Analytics team, partnering with data scientists and ML engineers to build secure, scalable data products powering AI-driven portfolio management.

- Built performant data pipelines using Python, SQL, Databricks and PySpark, enabling robust, reliable access to clean data for downstream modelling.
- Supported Chubb's cloud migration and implementation of Databricks Unity Catalog, transitioning legacy on-prem systems to Azure-native architecture.
- Designed and automated key workflows using Kafka and Databricks Workflows, improving efficiency and analytics throughput.

QUILTER FINANCIAL PLANNING | *DATA ENGINEER (CONTRACTOR)*

OCTOBER 2022 – NOVEMBER 2024 (2 YEARS)

Migrated on-prem databases to Azure Cloud, building robust pipelines with Azure Data Factory, Databricks, Python, SQL, and PySpark to ensure data quality and reliability. Worked in agile teams, contributing to scrums, stand-ups, and stakeholder alignment to deliver high-impact, business-driven solutions.

- Tech stack: Databricks, ADF, SQL, Python, PySpark, Tableau, Power BI
- DevOps: Git, Azure DevOps, Automated Testing
- Workflow: Agile, Scrum, Stakeholder Engagement, JIRA

WALLCACHE.COM | *FREELANCE*

FEBRUARY 2019 - MAY 2022 (3 YEARS)

Wallcache envelopes my freelancing exploits providing graphic design, portraiture and product photography, motion graphics, logo design, and creative consultation to clients. My website is: www.wallcache.com.

Education

2:1 BSC IN THEORETICAL PHYSICS, *IMPERIAL COLLEGE LONDON*

OCTOBER 2017 - JUNE 2020

Core Modules: Statistical Mechanics, Computer Science, Atomic Physics, Electromagnetism, Mathematical Methods, Fourier Analysis, Vector Calculus, Differential Equations, Medical Imaging & MRI, Nuclear Physics, Optics, and Quantum Mechanics. All modules contain strong elements of statistical analysis, experimental research, & programming.
Final year BSc Thesis: Conducted an in-depth mathematical derivation of the Black-Scholes-Merton model from physical first principles for predicting stock option prices, with the supervision of Professor Dimitri Vvedensky.
Year 2 Computing Project: Working with Object-Oriented Programming to investigate the performance of various lenses through Optical Ray Tracing.

BUSINESS ANALYTICS, *IMPERIAL COLLEGE BUSINESS SCHOOL*

FEBUARY 2022 - JUNE 2022

Used machine learning techniques such as regression analysis, nearest neighbour methods, decision trees, support vector machines and clustering techniques in Python. Summarised and visualised data to inform business decisions using linear programming and integer programming.

A-LEVELS & GCSE's, *MALVERN COLLEGE*

FEBUARY 2022 - JUNE 2022

A-Levels (2017): 3 A*s (Physics, Mathematics & Further Mathematics)
iGCSEs (2015): 13 A*s

Key Skills

- **Python Libraries**: Venv, Pandas, Numpy, BS4, requests, Data Science Pipelines, Matplotlib, NumPy, SciPy, Seaborn, Scikit-learn
- **Microservices Skills:** Flask API's and containerised cloud deployment
- **Cloud & Big Data Skills:** Cloud Data Lake, Hadoop architecture, Parquet SerDe, Hive and introduction to Snowflake
- **Database & Analytics:** Logical Data Modelling (Dimensional & 3NF), Advanced SQL, Data Visualisation in PowerBI & Tableau, Terraform.
- **Data Processing:** Spark DataFrames, Spark SQL, Databricks, Kafka, Azure Data Factory
- **SCM & DevOps**: Environments, Git, CI/CD, Docker, Automated Testing
- **Project Skills:** Stakeholder Engagement, Agile Workflow, JIRA, Scrum Master, Presenting