

Jack Grimes

jack@jackgrimes.net

London, UK

jackgrimes.net

GitHub: [jackgrimes](#)

LinkedIn: [jacketgrimes](#)

Google Scholar: [HZLogpQAAAAJ](#)

EDUCATION

IMPERIAL COLLEGE LONDON

PH.D. ENGINEERING/EPIDEMIOLOGY

Thesis: An investigation into the roles of water, sanitation, and hygiene in the control of schistosomes and other helminths

March 2012 - September 2015

IMPERIAL COLLEGE LONDON

M.Sc. ENVIRONMENTAL

ENGINEERING AND SUSTAINABLE DEVELOPMENT

September 2010 - September 2011

UNIVERSITY OF OXFORD

M.A. (OXON) PHYSICS

September 2007 - September 2010

SKILLS

PROGRAMMING

Very familiar: Python • R • Git/GitHub
Matlab/Octave

Familiar: SQL • JavaScript • HTML
CSS • \LaTeX

PLATFORMS

Google Cloud Platform • Domino

COURSES

Bayesian Statistics: From Concept to Data Analysis (Coursera, University of California, Santa Cruz)

Deep Learning Specialization (Coursera, deeplearning.ai)

Machine Learning (Coursera, Stanford University)

INTERESTS

Cycling • Walking • Travelling • Reading

EXTENDED CV

A more detailed version of my CV can be found on my [website](#)

EXPERIENCE

MORSUM London, UK

September 2022 - Present: **TECHNICAL LEAD (DATA SCIENCE)**

- Leading a squad responsible for product search, recommendations, and substitutions, in an online food marketplace ([.Anchovi](#))
- Productionising [Sentence-BERT](#) model serving and training
- MLOps: [Vertex AI](#) for model serving, data versioning
- Recruitment of a Machine Learning Engineer

March 2021 - August 2022: **SENIOR DATA SCIENTIST**

- Standardisation of incoming text data: the display name cleaner
 - Using text vectorisers and classifiers to fill out abbreviations, and to separate out product descriptions from names
 - Embedding solutions within wider business logic code, and productionisation on a Flask API
 - Ownership of the product and coordination with end users and other stakeholders
- MLOps: implementing an initial standardised approach to the model registry, loading and serving, using Google Cloud Storage
- Driving good software engineering practices: unit tests, code linting and formatting, and sonarcloud and GitHub Actions builds
- Building the Data Science team and outlining the teams' responsibilities
- Design of databases and pipelines for data ingestion: redefining data entities
- Git, Google Cloud Platform, MySQL, pandas, PyCharm, pytest, scikit-learn

LLOYDS BANKING GROUP London, UK

January 2019 - March 2021: **DATA SCIENTIST, APPLIED SCIENCE**

- Efficient, readable, unit-tested data science code for production
- Classification, clustering, network analysis, NLP: identification of incorrectly applied account and customer indicators
- Running the *Modelling Interest Group*, organising regular talks
- *Data Science Graduate Scheme Training Programme*: delivering the *Advanced Modelling* and *Natural Language Processing* sections
- Git, Jenkins, NetworkX, pandas, PyCharm, pytest, scikit-learn, spaCy

August 2017 - December 2018: **DATA SCIENTIST, RISK INNOVATION**

- Predictive modelling of fraudulent transactions, and of financial difficulty
- Topic modelling of regulatory data
- Face recognition for fraud prevention
- Presenting at internal conferences throughout the UK
- Data science solutions packaged into Flask apps for other teams
- face_recognition, Flask, Gensim, Git, pandas, SAS VA, scikit-learn

IMPERIAL COLLEGE LONDON London, UK

January 2016 - July 2017: **POSTDOCTORAL RESEARCHER**

- Investigating the environmental control of tropical parasitic diseases
- Mixed effects logistic regression (R and lme4) to compare environmental risk factors, parasites, and overall health, in a clustered dataset ([paper](#))
- Box-Cox transformations and linear regression (in R) to account for the 'long tail' of light parasite infections missed due to insufficient diagnostic sensitivity