

SUMMARY

Data Scientist and Engineer with expertise in geospatial data and imaging and experience in the end-to-end lifecycle of machine learning products from exploratory analysis through to deployment. Hold a **PhD in Bioinformatics**, and a **MSc in Computer Science**, with over 10 years programming experience. Worked with the *World Economic Forum*, *Netherlands' Transportation Ministry*, and the *London Fire Brigade*.

PROFESSIONAL EXPERIENCE

Numerico Technologies

Data Engineer

Feb 2018 – present

Data Scientist and Engineer

- Researched an inspector deployment that maximises network coverage on dutch highways, then assisted in productionising model building back-end Python services interacting with Kafka.
- Principal data scientist for a variety of research projects, for example using camera data to monitor traffic flow and speed, or analysing accident blackspots on highways.

Data Science for Social Good (DSSG)

textslData Scientist and Mentor

Jun 2017 – Jan 2018

DSSG Europe 2017 member and DSSG 2018 technical mentor.

- Developed an open-source fishing risk tool combining vessel tracking data with satellite imagery, using a random forest to score vessels according to likelihood of illegal fishing behaviours. Project code is available here: https://github.com/DSSG2017/wef_oceans

ASI Data Science / London Fire Brigade

Data Science Consultant

Jan 2017 – Apr 2017

Fellowship programme in commercial data science for researchers with strong analytic background.

- In house consultant with the London Fire Brigade, implementing topic modelling methods to classify corpus of 37,000 fire incident reports. Revealed previously unknown incident types, and visualised results. Project code: https://github.com/williamgrimes/london_fire_brigade

Laboratory for Molecular Cell Biology (LMCB)

PhD Student

Sep 2013 – Nov 2016

Joint scholarship at LMCB and A*STAR Bioinformatics Institute, Singapore.

- Applied machine learning techniques and segmented endothelial cell phenotypes in high-throughput microscopy assays. Analysed over 5TB of data with a detection accuracy of 82%.

National Institute of Informatics (国立情報学研究所)

International Internship Programme

Jun 2013 – Sep 2016

- Created annotation software in Java to assist classifying phenotypes in μ CT images of mice

EDUCATION

University College London

PhD Bioinformatics

Sep 2013 – Nov 2016

- Thesis: *Image processing and analysis methods in quantitative endothelial cell biology*

University College London

MSc Computer Science

Sep 2012 – Sep 2013

- Awarded Aardvark Scholarship 2012

Durham University

BSc Natural Sciences – Physics and Geophysics

Sep 2009 – Sep 2012

SKILLS / KEYWORDS

Programming: Python, R, Java, Matlab, SQL, Solidity, Bash and utils (sed, awk, grep)

Machine Learning: regression, decision trees, ensembles, SVM, neural networks, deep learning

Libraries: NumPy, pandas, SciPy, scikit-learn, scikit-image, OpenCV, pytorch, tensorflow

Databases: MySQL, PostgreSQL, MS SQL Server

Front-end: Javascript, HTML, CSS, Semantic UI, Jekyll

Misc: Amazon Web Services, Azure, Docker, Git, Kafka, Jupyter Notebooks, L^AT_EX, UNIX