

Ed Chalstrey

✉ edwardchalstrey@gmail.com | 🌐 edchalstrey.com | [in LinkedIn](#) | [🐙 GitHub](#)

PROFILE

Research software engineer with expertise in designing, deploying and maintaining secure cloud infrastructure, databases, full-stack applications, data visualisations and more. Highly experienced in collaborative working with interdisciplinary teams of scientists, software engineers and other diverse stakeholders. Able to communicate technical expertise to colleagues and clients with varied professional backgrounds.

EXPERIENCE

The Alan Turing Institute

Jan. 2019 – Present

Research Software Engineer / Research Data Scientist

London, UK

- Maintained and modernised Gambit, a C++ and **Python** library for game theory, implementing **CI/CD** pipelines with **GitHub Actions**. Developed a dataset for **generative AI (LLM)** code generation via natural language game descriptions and an associated **React/Flask** web application integrating live LLM calls. Improved interoperability of Gambit with OpenSpiel for **reinforcement learning** experiments.
- Developed Autoemulate, translating sophisticated **PyTorch** machine learning models into a robust Python package. Designed a low-code API that abstracted complex technical details, enabling research scientists to easily deploy AI-accelerated emulator models for physical simulations.
- Built scalable data pipelines to engineer a harmonised UK solar PV **PostgreSQL** database from diverse, unstructured sources. This infrastructure formed the basis of a dataset published in *Nature Scientific Data*, directly enabling subsequent AI-driven solar nowcasting models.
- **Data wrangling** and database creation/curation for researchers accessing electronic healthcare records for ML projects in Trusted Research Environments (TREs).
- Designed, deployed and managed instances of **Data Safe Haven**, a TRE infrastructure-as-code tool, using **Pulumi** for IaC, **Docker** for containerization, and **Azure** cloud services.
- Developed the Seshat Global History Databank, a full-stack web application using **Django** and **PostgreSQL** for data management and Leaflet.js for interactive geospatial visualization. Involved **API** and **GUI** design and user testing.
- Other projects included building the open-source Scivision computer-vision library, and authoring a technical report on ML usage for the UK Census.
- Mentored junior colleagues and taught software engineering/data science courses to PhD students.
- Managed stakeholder relationships varying from computer science professors at leading universities, external data provider organisations involved in Data Study Groups (cloud-hosted secure data hackathons) I facilitated, to internal and external software engineer collaborators and open-source software contributors.
- Academic publications including several in *Nature Scientific Data* and talks at conferences.
- Hosting the institute's AI themed podcast and Entrepreneurship interest group, inviting guest speakers and experience pitching ideas to VCs.

WallStreetDocs / Citibank

Sep. 2015 – Jan. 2019

Forward Deployed Engineer / Analyst

London, UK

- Developer role in a FinTech startup implementing changes and bug fixes in a document automation platform for investment bank clients using PHP.
- Worked as a consultant at Citibank (Oct 2016 - Jan 2019), handling high volumes of client demands and conflicting deadlines.

UCB

May 2017 – Aug. 2018

MSc Research Project Student

Slough, UK

- Developed software for *in silico* drug efficacy prediction based on drug target data.
- Developed a Python tool utilising the Open Targets Platform to enable searching of drug targets and retrieval of prioritised disease candidates.

University College London (UCL)

May 2015 – Aug. 2015

Research Intern

London, UK

- Extended the Django-based Python codebase for the OMA orthology browser to include domain architecture images for protein sequences.

The Sainsbury Laboratory

Research Intern

Aug. 2013 – Jul. 2014

Norwich, UK

- Developed a bioinformatics genetic algorithm to locate phenotype-causing SNPs in mutagenized plant genomes based on homozygous/heterozygous SNP density.

EDUCATION

Birkbeck, University of London

MSc Bioinformatics with Systems Biology (*Distinction*)

Oct. 2016 – Aug. 2018

London, UK

University of Essex

BSc Biological Sciences (2.1)

Oct. 2010 – Jun. 2013

Colchester, UK

SKILLS

Software Engineering : TDD, Agile, Technical Writing, Databases

Languages : Python, R, Ruby, JavaScript, TypeScript, SQL, HTML/CSS

Frameworks/Libraries : Django, FastAPI, React, Pandas, NumPy, PyTorch, SQLAlchemy, Pydantic

DevOps/Cloud : Docker, Pulumi, Azure and AWS, GitHub, CI/CD, Linux, HPC

PUBLICATIONS

[Cliopatria - A geospatial database of world-wide political entities from 3400BCE to 2024CE](#) *Nature Scientific Data* (2025).

[A harmonised, high-coverage, open dataset of solar photovoltaic installations in the UK](#) *Nature Scientific Data* (2020).

[Ten simple rules for working with other people's code.](#) *PLOS Comp. Bio.* (2023).

[Digital identity architectures: comparing goals and vulnerabilities](#) *arXiv* (2023).

Technical reports on [SATRE- a Standard Architecture for Trusted Research Environments \(TREs\)](#) (2023) and [TRE best practices](#) (2021).

Blog posts on [Using generative AI to write code: a guide for researchers](#) (2024) and

[Openness in the digital identity context](#) (2021).