Ed Chalstrey

edwardchalstrey@gmail.com
uk.linkedin.com/in/edchalstrey | github.com/edwardchalstrey1
DOB: 23/08/1992

Profile

Currently employed as a Research Data Scientist at The Alan Turing Institute. Background in computational biology research and experience handling clients for a SaaS company.

Skills and Experience

- Broad technical/soft skills via wide variety of software, data science and bioinformatics projects
- Designing and building scientific software with Python, R, Ruby, JS, SQL and more
- Analysis and visualisation of research data and statistics with Python and R
- Knowledge of best practices in software and data engineering, data governance
- Familiarity with common ML frameworks, the ML data lifecycle and MLOps
- Packaging of scientific code into well tested and maintainable packages
- Version control with Git/GitHub/GitLab including as part of large collaborative projects
- Use of GitHub for project management and issue tracking
- Containerisation (Docker) and working with Infrastructure as Code (Pulumi)
- Familiarity with Linux, virtual machines, use of high-performance computing infrastructure
- System administration of Linux cloud environments on MS Azure (TREs)
- Test-driven software development experience and CI/CD (pytest, GH actions)
- Data modeling, database design, data wrangling with PostgreSQL, Python, SQL alchemy
- Front end web development with e.g. JavaScript and ReactJS, Django
- UI/UX design and user testing
- Bioinformatics research and software development experience
- Analysis of NGS data and other biological data formats
- GIS development experience with tools such as Leaflet, Geo-Django
- Pair programming and code review with colleagues
- Maintenance of production code for both academic partners and in industry
- Contributions to open-source software packages
- Development of project documentation for academic research software and services
- Experience managing and deciding research project objectives and delegation
- Collaboration on interdisciplinary academic projects and communicating with diverse stakeholders
- Experience dealing with high volume of client emails/demands and conflicting deadlines
- Line management of junior colleagues
- Presentation of original research and software skills at conferences/seminars/talks
- Teaching Research Software Engineering and Data Science courses for PhD students
- Scientific report writing and contributions to academic research papers
- Scientific workshop and event management and facilitation

Employment

Dates	***Company***		**Job title/info**
12/20 - present	The Alan Turing Institute		*Research Data Scientist*
12/19 - 12/20	The Alan Turing Institute		*Junior Research Data Scientist*
01/19 - 12/19	The Alan Turing Institute		*Research Assistant*
05/17 - 08/18	UCB		*Project student* – MSc Research project. Research info below.
10/16 - 01/19	Citi		*Consultant from WallStreetDocs*

09/15 - 01/19	WallStreetDocs		*Analyst* – PHP coding role, working on a document automation platform with investment bank clients.
05/15 - 08/15	University College London (UCL)		*Research Intern* - 3 month computational biology internship. Research info below.
08/13 - 07/14	The Sainsbury Laboratory (TSL)		*Research Intern* - 12-month graduate internship in computational biology. Research info below.

Education

Dates	School	Course	Grade
10/16 - 08/18	Birkbeck, University of London	*MSc Bioinformatics with Systems Biology*	*Distinction*
10/10 - 06/13	University of Essex	*BSc Biological Sciences*	*2.1*

Research

The Alan Turing Institute:
Variety of research software engineering, data science and infrastructure development projects across academic disciplines ranging from environmental science, healthcare data science, digital humanities and more. Projects include but not limited to; developing an in-house service for data science projects on sensitive data in Trusted Research Environments, updating/enhancing existing code bases for academic project software ranging from tackling misinformation to crowdsourcing urban environment data, creating an interactive data visualisation for pollution spread in water networks, drafting a technical report to scope a future machine learning project for the UK census, building an open source computer vision python library and contributor community.

Birkbeck, UoL MSc project with pharmaceutical company: UCB
Development of software for *in silico* drug efficacy prediction, based on drug target data. Utilizing the free online software “Open Targets Platform” which compiles/validates evidence for target-disease associations from multiple data sources and evidence types, I developed a new tool to enable the searching of a list of drug targets and retrieval of prioritised disease candidates.

UCL Summer Internship Project:
Computational biology project that involved extension of the Django-based Python code for the OMA orthology browser, to include domain architecture images for the protein sequences.

TSL 12-month Internship Project:
Bioinformatics algorithm development project, with an aim to locate phenotype-causing SNPs in backcrossed mutagenized plant genomes. A genetic algorithm approach was used to find the correct permutation of unordered sequence contigs, based on the expectation of a peak in the ratio of homozygous to heterozygous SNP density, around the location of the causative SNP.

University of Essex Dissertation Project in Molecular Science:
Bioinformatics research project involving the differential profiling of transcriptomes from stage 1a lung adenocarcinoma tissue and adjacent healthy lung cells. Gene Ontology annotation and enrichment analysis were used to identify potential targets for interfering drug therapy.

Other

I have gained additional experience communicating complex scientific topics to general audience via The Turing Podcast