

Results-driven team player with extensive expertise in Python development, machine learning, GIS, and bioinformatics. A passionate problem solver who excels in bridging the gap between data science and software engineering. Experienced in working on groundbreaking projects and delivering innovative solutions.

## Experience

---

### **Cell Therapy & Bioinformatics Consultant (Contract) at Allogene**

November 2021 - Present

Collaborating with Allogene on a consulting basis to develop an advanced off-target detection pipeline for the groundbreaking [AlloCar T Therapy](#), an innovative cell therapy for cancer treatment. Integrating Chip-Seq data and in-silico TALEN binding activity modeling, we assess the probability of true-positive and false-positive events within a terabyte-scale data pipeline. Established a stringent approach to guarantee patient safety and provide compelling evidence of safety to regulatory bodies, including the FDA, considering the critical nature of the decisions made.

Technologies: Python, Google Cloud Platform, data pipeline tools, bioinformatics tools.

### **Data Science Consultant at HSBC**

September 2022 - June 2023, London

Collaborated with the financial fraud detection team to provide expert guidance on model deployment, productionisation, and testing strategies. Spearheaded the creation of an internal dashboard that elevated decision-making processes, empowering stakeholders with real-time, intuitive, and actionable insights into model performance. Put processes in place for the team to deploy and test models within a highly regulated environment, while ensuring key performance metrics.

Technologies: Python, Kubernetes, BigQuery, SQL, Pandas, Streamlit, Model development, testing, and productionisation

### **Senior Applied Scientist at Improbable**

October 2020 - August 2022, London

Working as part of the defence team to simulate civilian-military interaction using agent-based modelling. Helped to implement best practice Python methodology to upskill the team and substantially increase team productivity. Developed significant expertise in geospatial analysis, visualisation, and its use in population modeling.

Technologies: Python, Google Cloud Platform, Kubernetes, agent-based modelling, geospatial analysis

### **Python/Machine Learning Engineer (Contract) at Deloitte**

October 2019 - July 2020, London

Developed simulation engine for Network Rail to assist incident recovery planning. Developed prescriptive AI route planning models to advise on the best course of action following incidents on the track.

Technologies: Python, Google Cloud, Kubernetes, Big Query, Neo4j, Docker, Graph algorithm analysis & development

### **Vice President of Data Science at Celixir**

June 2018 - August 2019, London

Led team of developers through a wide range of projects including a deep learning pipeline to predict cell health based on image data, and a small molecule generation platform used by lab scientists that cross-referenced patent information to avoid competing IP.

### **Chief Technology Officer at Desktop Genetics**

April 2018 - August 2019, London

Led a highly successful team of bioinformaticians and software engineers. Spearheaded development and automation of TheraSeq - a clinical genome analysis pipeline operating on terabytes of patient data. Reduced required employee hands-on time per TheraSeq project from 3-4 months to 2 days. This assisted in a 63% increase in average gross margin percentage per project. TheraSeq has since been used in the two successful FDA drug filings. Developed novel in-silico testing and benchmarking techniques.

Remained active part of the data science team, developing machine learning models to predict biological processes, and TIDE - a tool used by hundreds of scientists every day to analyse their lab results.

### **Data Scientist at Desktop Genetics**

October 2016 - April 2018, London

Worked as part of the R&D team working on problems in the genetic engineering field and specifically with CRISPR. Projects included modeling CRISPR on-target and off-target activity, and designing experiments to maximize information utility. Developed novel CRISPR on-target activity scoring function (Dunne 2017) that out performed the state of the art. Presented results to clients, and at Cold Spring Harbor Laboratory. Assisted clients in the design and analysis of CRISPR genetic profiling and drug discovery experiments.

Technologies: Python, Pandas, Scikit-Learn, PyMC, TensorFlow, PostgreSQL, Dask, Apache Spark, Google Cloud, Ansible

## Analyst at Goldman Sachs

July 2015 - October 2016, London

Responsible for developing and maintaining software used daily by the Interest Rates sales team. Interacted daily with traders and sales people to support the software and understand the key requirements. Involved in a natural language processing project to rapidly process trade inquiries, saving the company large amounts of otherwise lost revenue.

Technologies: Java, Scala, C#, Slang, Javascript, Python, NLP, Apache Kafka, RabbitMQ

## Summer Analyst at Goldman Sachs

Summer 2014, London

Developed a key piece of in-house Interest Rates trading software used to process billions of dollars of transactions daily. Full stack development included everything from provisioning the bare-metal machines for the back end to working with end users to design the UI on the front end. Placed as campus ambassador at University College Cork for Goldman Sachs following exceptional performance.

Technologies: C#, Java, Slang, Javascript

## Intern Software Engineer at HeyZap

Summer 2013, San Francisco

Developed innovative mobile adverts using HTML5 and CSS3. Quickly became the go-to guy for CSS related issues.

## Education

---

### University College Cork

2011 - 2015

B.Sc (Hons) Computer Science, First Class Honours

**Key Subjects:** AI, Machine Learning, Algorithm Analysis, Theory of Computation, Linear & Abstract Algebra

#### Awards:

- *Title of College Scholar* - Awarded in 2013, 2014, and 2015 for achieving First Class Honours in end of year exams.
- *Most Innovative Final Year Project* - Awarded by Fidelity Investments for work on the application of Machine Learning to the stock market. Analysed market prediction techniques and robustly evaluated their predictive value.

#### Online Courses

- *edX 2016 - Introduction to Biology - The Secret of Life*
- *Coursera 2015 - Machine Learning*
- *Coursera 2014 - Cryptography I*
- *Coursera 2013 - Computational Investing*

### Ballincollig Community School

2005 - 2011

**Key Subjects:** Higher Level Mathematics (A2), Higher Level Applied Mathematics (A2), Higher Level Physics (B1)

Represented school in mathematics and linguistics competitions

## Key Skills

---

- **Management:** Personable, pragmatic, and goal orientated. Track record of meeting ambitious team goals.
- **Data Science:** Breadth of experience in machine learning and statistical modeling. Expert in PyData Stack (**Pandas**, **Numpy**, **Scikit-Learn**, **SciPy**, **Matplotlib**, etc). Experience using **Tensorflow** to build high performance deep learning models, and **PyMC** to perform Bayesian statistical modeling. Ability to communicate technical statistical results to a non-technical audience.
- **Data Engineering:** Developed pipelines to process and model terabyte sized datasets. Experience using **Apache Spark** and **Dask** to run distributed machine learning pipelines over **Google Cloud** clusters. Familiar with **Apache Kafka** and **RabbitMQ**. Experience with **PostgreSQL**, **NoSQL** databases, **Apache Parquet**, **BigQuery**.
- **Software Engineering:** Experience in working with enormous codebases (over 100M LoC), and within teams. Strong understanding of how to write high quality maintainable code. Experience using **Git**, **SVN**, and **CVS** version control systems.
- **GIS:** Hobbyist and sometimes occasionally professional map maker. Good understanding of GIS ecosystem and working with spatial data.
- **Bioinformatics:** Good understanding of genomics, focusing on cell and gene therapies including CRISPR and TALEN. Familiar with the ecosystem of bioinformatics tools with a focus on variant detection and analysis.
- **Programming:** Experience with **Python (10+ years)**, **Java**, **Scala**, **Ruby**, **R**, **C#**, **Haskell**, **Javascript**, **Slang**, **HTML/CSS**.