# Owais Siddiqi

omsiddiqi01@gmail.com | owaissiddiqi.co.uk | London, United Kingdom

## Profile

I am a data scientist specialising in **bioinformatics**, **data analytics**, and **multi-omics**, with proven expertise in **statistical modelling**, **clinical data analytics**, and large-scale multi-omic dataset analysis. My experience includes developing and optimising automated **bioinformatics workflows**, integrating **AI-driven analytics** and **computational biology** to enhance accuracy, efficiency, and reproducibility in biomarker discovery and **precision oncology**. I excel in collaborative, cross-disciplinary teams and continuously advance my skills to remain at the cutting edge of genomics research and AI-driven healthcare innovation.

## Education

Imperial College London *Oct. 2022 – Oct. 2023*
MSc in **Biomedical Engineering**

- Developed expertise in **Reinforcement Learning** (DQN, PPO, SAC), **Bayesian Probability**, and **Markov Decision Processes**.
- Applied **deep learning** models for image classification, object detection, and segmentation.
- Conducted advanced **time series analysis**, neural decoding, and **machine learning** for **brain-machine interfaces**.

Queen Mary University of London *Sept. 2019 – Sept. 2022*
BEng (Hons) in **Biomedical Engineering** – 1st Class

- Mastered calculus, differential equations, eigenvalues, vectors, and Gaussian elimination for engineering applications.
- Strengthened skills in multivariable calculus, linear algebra, and vector analysis for biomedical signal processing.

## Work Experience

**ANGLE plc.** *Jan. 2025 – Present*
*R&D Data Analyst*

- **Clinical and Bioinformatics Data Management:** Analyse large-scale imaging and multi-omic datasets for biomarker discovery in liquid biopsy and circulating tumour cell (CTC) applications. Implement automated, reproducible bioinformatics pipelines for sequencing data analysis, enhancing scalability and consistency.
- **Workflow Automation and Optimisation:** Develop Python and Bash scripts automating data extraction, transformation, analysis, and reporting, significantly reducing manual workload.
- **Statistical and Computational Analysis:** Apply statistical methods and machine learning approaches to validate experimental data, optimise research design, and enhance biomarker selection for precision oncology.
- **Cross-Functional Collaboration:** Work closely with scientists and IT teams to integrate data solutions and ensure regulatory compliance with ISO 13485:2016+A11:2021 standards.
- **Precision Oncology Advancement:** Leverage multi-omic technologies to advance biomarker identification, accelerate drug development, and improve clinical decision-making.

**Genevation Ltd.** *Apr. 2024 – Dec. 2024*

*Junior Genomic Data Scientist*

- **Cancer Biomarker Prediction:** Developed AI models using PyTorch for neoantigen biomarker prediction, facilitating personalised vaccine production for cancer immunotherapy.
- **Bespoke VCF Production:** Designed a specialised pipeline converting raw FASTQ data into customised VCF formats, enhancing downstream genetic analyses of tumour-specific biomarkers.
- **Bioinformatics and Computational Tools:** Processed RNA-seq data and performed variant calling using Docker containers and managed Linux-based virtual machines on Google Cloud Platform.
- **Optimised Data Processing:** Accelerated genetic downstream analysis of neoantigen biomarkers from one week to a few days.
- **Continuous Learning:** Regularly participated in workshops and online courses to stay current in genomics, AI, and bioinformatics advancements.

**PDUK Ltd.** *Aug. 2023 – Jan. 2024*

*Data Scientist Intern*

- **Data Engineering:** Utilised SQL queries in MySQL for data cleaning and management, reducing data processing time by 25% across 3000+ customer records.
- **Data Visualisation:** Created visualisations using Pandas, Matplotlib, and Seaborn, improving insight and decision-making through clear, data-driven reports.
- **Predictive Modelling:** Leveraged recurrent neural network (RNN) models in PyTorch for predictive time-series forecasting, improving accuracy by 20% and reducing inventory costs by 10%.
- **Customer Segmentation:** Implemented K-means clustering for customer segmentation, resulting in a 15% increase in targeted marketing conversion rates.

**Imperial College London (MSc Project)** *Mar. 2023 – Sept. 2023*

*AI Researcher (Deep Learning)*

- **Bespoke Model Development:** Created a novel regression AI model using microscopy data for cell differentiation, achieving over 80% accuracy.
- **Data Management and Augmentation:** Implemented advanced data preprocessing and augmentation methods, improving model generalisation by 30% and improved predictive accuracy.
- **Tool Integration:** Integrated Pandas, Matplotlib, OpenCV, TensorFlow, Keras, and NumPy, optimising data preprocessing, parameter extraction, and model implementation, ensuring an efficient and effective workflow.

## Technical Skills

- **Programming:** Python, MySQL, Bash, MATLAB
- **Machine Learning:** TensorFlow, PyTorch, Scikit-learn, Keras
- **Bioinformatics:** Genomics, Transcriptomics, FASTQ–VCF Pipelines, Nextflow
- **Cloud/DevOps:** Google Cloud, Docker, Git, CI/CD, VMs, Linux
- **Data Analysis:** Statistics, Statistical Modelling, Visualisations

## Publications

**Siddiqi, O.**, Winfield, J., Sormpas-Petridis, K., Harris, E., Ramkumar, A., Candito, A., Görner, S., Messiou, C., Blackledge, M., & Thrussell, I. (2022). *Investigating the Repeatability of Multifrequency Magnetic Resonance Elastography applied to a Soft Gelatine Phantom.* Presented at the ISMRM 31st Conference & Exhibition, London, 2022.