

# Ghulam M. Khan

Email: gmkghulam2000@gmail.com | Mobile: +44 7946 597411 | GitHub: <https://github.com/GMK2000GMK>  
LinkedIn: <https://www.linkedin.com/in/ghulam-muhammed-khan>

## EDUCATION

### Reading, University of Reading

Sep 2019 – Jul 2022

BSc (Hons) Biomedical Science (Grade 2:1)

Key modules: Biochemistry, Bioinformatics, Clinical skills, Genomics, Laboratory, Neuroscience, Infectious Diseases.

Final year dissertation paper on 'The effects of diet on the gut microbiota and neurological diseases'.

### London University of Birkbeck

Sep 2022 – Apr 2024

MSc (Hons) Bioinformatics (Grade Distinction)

Key modules: Biocomputing, Omics, Statistics, Structural Biology, Molecular Biology.

Completed thesis: "The effect of read length and nanopore base calling quality on metagenomic classifier results".

Leaving Cert: 2 successful H2(80%-89%) in Mathematics, and Biology

## PROFESSIONAL EXPERIENCE

### Bioinformatician (2 year Contract), NHS/University of Oxford

May 2022 – May 2024

- **Write and Automate Workflows:** Utilise Python, R, bash scripting and LLM-based solutions to automate workflows involving basecalling, quality control, alignment, variant detections and visualisation, to streamline processes across various projects while generating natural language summaries for clinical reports.
- **Problem Solving and Collaboration:** Collaborate with cross-functional teams to analyse complex problems, recommend technology-driven solutions, and devise innovative software strategies using Slack, involving biostatisticians, lab technicians and other stakeholders.
- **Application Development:** Develop and customise software applications using AWS, Docker, Github, and Node.js to meet specific project requirements, focusing on scalability and performance.
- **Data Analysis:** Compile and analyse data using semantic web ontologies, Power BI and Tableau for reports and Python libraries such as Heapq, Pandas and NumPy for data manipulation, in synergy with Matplotlib and Seaborn for visualisation.
- **Time Management and Multitasking:** Successfully balanced employment at the University of Oxford, managing work-related projects alongside MSc coursework, projects, and exam preparation.
- **Communication:** Communicate project findings and software solutions through presentations, technical documentation, and comprehensive project reports using Microsoft office, Power BI and Tableau.

### Employment Gap

May 2024 – Dec 2024

- **Travel:** Engaged in personal travel and cultural exploration.
- **Personal development:** Dedicated time to advancing my knowledge in Ontology, HTML, Node.js, and Java through self-directed learning. Completed various coding challenge projects as part of a personal passion for programming and skill development.
- **Job search**

## SKILLS

**Technical skills:** AWS, Oracle, Python, R, SQL, Bash, JavaScript, HTML, Node.js, CSS, MongoDB, Apache, Linux, Docker, Kubernetes, Ansible, LLM, NLP, NGS.

**Software and Packages:** Abbricate, Apache Jena, Ansible, Bcftools, Bracken, BLAST, BWA, Centrifuge, Clair3, Diamond, Django, Docker, Dragonflye, FastQC, Flask, GATK, Github, Gubbins, Heapq, Java, Jenkins, Kaiju, Kraken, Mafft, Minimap2, MLST, Nanoplot, Nanostat, Nextflow, Numpy, Next Generation Sequencing (NGS), OWL, Pandas, PostgreSQL, Pymol, Pysamstats, Quast, SAMtools, SHACL, Shovill, Snippy, Unicycler, Slack.

**Programming Languages:** Bash, HTML, Node.js, Perl, Python, R, SQL, C++, SPARQL, Unix.

**Machine Learning:** BioPython, Keras, Pytorch, Scikit-learn, TensorFlow, XGBoost

**Data Visualisation and Reporting:** Excel, Microsoft Power BI, Protege, PowerPoint, RDF Format, Tableau, Word.

**Hardware:** Cloud, Datacentres, Illumina, Infrastructure, Nanopore, Networking, Servers.

**Soft skills:** Adaptability, analytical thinking, communication, problem-solving, teamwork

**Languages:** Fluent in English, Conversational in Spanish

## MSc Bioinformatic Project :

- Investigated the impact of nanopore read length and basecalling quality on metagenomic classifier accuracy.
- Identified optimal parameters to improve taxonomic classification sensitivity and precision.
- Enhanced reliability of metagenomic analysis for microbial community studies

### Evaluating DNA Preservation with patented methodology:

- Conducted a comprehensive evaluation of DNA preservation methods using Nextflow pipelines with tools such as Biopython, Minimap2, Nextflow, VCFtools, Samtools, etc.
- Compared Nanopore and Illumina sequencing technologies to assess key metrics such as depth and coverage.
- Identified significant advancements in DNA preservation and sequencing accuracy.
- Worked closely with wet lab biologist, biostatisticians and other bioinformaticians.

## CERTIFICATIONS

- AWS Certified Cloud Practitioner Course
- Python A-Z™: Python For Data Science With Real Exercises
- R Programming A-Z™: R For Data Science With Real Exercises
- Data Science A-Z: Hands-On Exercises
- Code Institute HTML & Javascript Coding Challenge
- Practical Knowledge Modelling Ontology Development 101