

# QIANRONG GUO

+44 7536-128-315 | [chloeguo830@gmail.com](mailto:chloeguo830@gmail.com) |  
[linkedin.com/in/qianrong-guo](https://www.linkedin.com/in/qianrong-guo) | [github.com/Rong830](https://github.com/Rong830)

## EDUCATION

- Doctor of Philosophy** | *Artificial Intelligence for Phenotypic Virtual Screening* Jan. 2023 – Jan. 2027  
Imperial College London London, England
- Master of Science** | *Major: Bioinformatics (With Distinction)* Sep. 2021 – Aug. 2022  
The University of Edinburgh Edinburgh, Scotland
- Related courses: Machine Learning (A); Natural Language Processing (A); Bioinformatics Algorithms; Statistics
- Bachelor of Science** | *Major: Biological Science* Aug. 2017 – May. 2021  
China University of Geosciences Wuhan, China
- Related courses: C Language Programming (A+), Mathematical Statistics (A), Bioinformatics (A), R Language and Statistic (A-)

## RESEARCH EXPERIENCE

- AI for Phenotypic Drug Discovery against Glioblastoma Cells** Mar. 2022 – Apr. 2022  
The University of Edinburgh Edinburgh, Scotland
- **Aim:** Machine learning approaches for discovering novel chemicals that change the morphology of Glioblastoma cells. (**Project Report**).
  - **Method:** Developed and optimized a graph neural network for feature extraction combining information from RDkit and features extracted from a large pre-trained language model (ChemBERTa). Improved model architecture and metrics, particularly for imbalanced data.
- Effects of red phosphorus on metabolism of microorganisms** Mar. 2020 – Jun. 2021  
China University of Geosciences Wuhan, China
- **Aim:** Identification of novel chemicals that interfere with the metabolism of microorganisms.
  - **Method:** Applied Gradient Boosting Tree and Random Forest to predict the most efficient concentration of 4 types of red phosphorus using transcriptomics data and visible light wavelength. Performed data cleaning, analysis, and visualization using Python.
- Botany and Ecology Feature Analysis** Jul. 2019 – Sep. 2019  
China University of Geosciences Zigui, China
- **Aim:** Evaluation of the forested landscape of the Kuilongshan Nature Reserve using qualitative analysis and landscape evaluation index systems.
  - **Achievement:** Analyzed main features of the poaceae family in mountainous environments. Conducted data collection, analysis, and visualization using MATLAB. Wrote report.
- Neosinocalamus Growth Prediction Using Image Data** Mar. 2019 – Jul. 2019  
China University of Geosciences Wuhan, China
- **Aim:** Analysis of the connection between the growth and development of Neosinocalamus affinis and the cells' structure and components at the molecular level.
  - **Method:** Collected and processed leaves from Neosinocalamus affinis. Applied k-means clustering to the mean size and number of phytoliths from different leaf positions and ages. Discovered the correlation between phytolith size and number and plant growth, which can be used to predict age and living environment of plant samples collected from fossils.

## WORK EXPERIENCE

- Development Intern** Sep. 2022 – Dec. 2022  
WuXi Biologics - Full-time Shanghai, China
- **Department:** Data Technology and Analysis Department - Data Technology AI Team
  - **Responsibilities:** Developed graph-based and language models for protein-protein interaction (PPI) prediction and antibiotic discovery.
  - **Skills:** Graph-based and language model development, protein-protein interaction prediction, antibiotic discovery.

## SELECTED PROJECTS

---

**Implemented Neural Machine Translation using the Transformer architecture** Feb. 2022 – Apr. 2022  
The University of Edinburgh Edinburgh, Scotland

- **Goal:** Train an NMT model for German to English and evaluate its performance with BLEU score.
- **Contributions:** Modified the decoder to do beam search and implemented the lexical model using LSTM. Implemented linear projection of query, key, and value, scaled dot-product attention for h attention heads, and concatenation of heads and outputted projection.
- **Skills:** Neural machine translation, transformer architecture, BLEU score evaluation, beam search, LSTM, linear projection, attention heads.

**Web Site Design for Molecule Compounds using HTML and PHP** Jan. 2022 – Apr. 2022  
The University of Edinburgh Edinburgh, Scotland

- **Goal:** Build a database and web pages for a molecule storage application.
- **Contributions:** Developed web pages using style sheets and CSS and added animation using JavaScript.
- **Skills:** HTML, PHP, CSS, JavaScript.

**Building a Recurrent Neural Network Language Model from Scratch** Jan. 2022 – Feb. 2022  
The University of Edinburgh Edinburgh, Scotland

- **Goal:** Train a recurrent neural network (RNN) to predict subject-verb agreement and number prediction with an RNN language model (RRNLM).
- **Contributions:** Implement basic functionality of an RNN for language modeling. Given a word sequence  $w_1, w_2, \dots, w_t$ , a language model predicts the next word  $w_{t+1}$  by modeling:  $P(w_{t+1}|w_1, \dots, w_t)$ . Implemented its core word prediction functionality and its training by implementing a loss function and the model's gradient accumulation through backpropagation. Trained and fine-tuned the model on actual data and used it to generate sentences.
- **Skills:** RNN, language modeling, loss function implementation, gradient accumulation through backpropagation, data fine-tuning.

**Protein Sequences Identification Program (Python)** Oct. 2021 – Dec. 2021  
The University of Edinburgh Edinburgh, Scotland

- **Goal:** Identify a family of protein sequences from a user-defined subset of the taxonomic tree and determine the level of protein sequence conservation across species within the group. Scan protein sequences with motifs from the PROSITE database to identify known motifs (domains).
- **Contributions:** Obtained user-defined protein sequences from NCBI and visualized obtained protein sequence distances using seaborn.
- **Skills:** Protein sequence identification, taxonomic tree analysis, protein sequence conservation analysis, motif identification.

**Differential Expression Analysis on Trypanosoma congolense Genome** Sep. 2021 – Nov. 2021  
The University of Edinburgh Edinburgh, Scotland

- **Goal:** Process RNA-Seq sequencing data and perform differential expression analysis on a high-performance computer (HPC).
- **Contributions:** Built an interactive program to perform alignment and expression analysis using bash. Analyzed the quality of fasta files using fastqc and aligned sequences using bowtie2, samtools, and bedtools.
- **Skills:** Differential expression analysis, RNA-Seq data processing, alignment, fasta file analysis, bash programming.

## AWARDS

---

**Scholarship of Outstanding Ability** 2017 - 2021  
Awarded to students with outstanding academic and leadership abilities.

## SKILLS

---

**Languages:** Chinese (Native), English (IELTS 7)

**Programming:** Python (NumPy, PyTorch, sklearn, Matplotlib, Pandas, seaborn), Shell, R, PHP, SQL, Java, JavaScript, C, MATLAB

**Document Creation:** LaTeX, Microsoft Office Suite, Markdown