

GABRYEL MASON-WILLIAMS

✉ gabryelmw@gmail.com

🌐 <https://gmw99.github.io>

🌐 [linkedin.com/in/gabryel-mason-williams/](https://www.linkedin.com/in/gabryel-mason-williams/)

📄 [Google Scholar](#)

Education

PhD in Artificial Intelligence

Queen Mary University of London

Sep. 2023 – Sep. 2026

Thesis Title: Artificial Neuroscience: Applying Mathematics to the Understanding and Control of Deep Learning

Awards: Principle Scholarship covering tuition fees, and provides annual tax-free maintenance for 3 years
Google 2023 Academic Scholarship of €7000

MSc in Artificial Intelligence

Queen Mary University of London

Distinction 85.32%

Sep. 2022 – Sep. 2023

Dissertation Topic: What Makes a Good Prune [1] (Published at ICLR)

Core Modules: Machine Learning, Neural Networks and Deep Learning, Artificial Intelligence, Applied Statistics

BSc in Computer Science

University of Plymouth

First Class 76.26%

Sep. 2017 – Sep. 2021

Dissertation Topic: Automating Lab Book Data Collection With Edge Machine Learning

Final Year Modules: Computing Project, Machine Learning, Parallel Computing, Alternative Paradigms, Advanced Computing and Networking Infrastructures

Awards: Deans Lists 2017/18

Technical Skills

Languages: Python, Bash, C

Frameworks: Scientific Python: (Jax, Numpy, Scipy, Matplotlib), Machine Learning: (Pytorch, Flax, Tensorflow), Distributed and Parallel Systems: (Tensorstore, Rados, Boto3, mpi4py, Ray)

Software Engineering: Version Control, Code Review, Documentation, Testing, Object Oriented Programming

Experience

Research Software Engineer

Jul 2021 – Present

Rosalind Franklin Institute: Artificial Intelligence & Informatics team

Remote

- Creator and sole developer of DisTRaX, which is a Python application for deploying temporary storage systems onto High-Performance Compute (HPC) infrastructure in a fast and scalable fashion. For our Cryo-EM software, it reduced processing times and I/O overhead by 4.37% and 100%, respectively.
- Lead developer for Kompressor, a machine learning research project to create a novel neural losses compression algorithm for large volumetric scientific data. Which aims to reduce data overhead and storage costs significantly.
- Involved a successful grant bid for £10,000 in research funding for "Using Novel Cluster Technology to Improve Performance of Cryo-EM Analysis" with the University of Bristol.
- Played an active role in the creation and teaching of the lessons in machine learning for the first-year PhD students.
- Effective dissemination of research via talks at conferences [2], [3], [4]

Maths Circle Mentor

Oct 2021 – Jul 2022

Exeter Mathematics School

Remote

- Mentored students in years 8 and 9 in mathematics to encourage the development of mathematical thinking outside of the school context.
- Successfully improved their ability to be adventurous, articulate and accepting by providing them with challenging questions and creating an effective collaborative learning environment

Scientific Computing Research Placement: Year in Industry Student

Jun 2019 – Sep 2020

Diamond Light Source

Oxford/Remote

- Researched "High-Performance Object Stores for Intermediate Data processing"
- Created GRAM [5] a kernel module to represent RAM as a block device for Ceph and DisTRaC [6] an application that allows Ceph to run on a High-Performance Computing cluster using RAM, which led to a reduction in I/O and processing times of 81.04% and 8.32%, respectively.
- Presented a poster at Ceph Day CERN and gave a talk/demo of DisTRaC to the Ceph Science Group.
- Created a proof of concept for the visualisation of objects inside an object store in real-time, which was presented at the SuperComputing 19 conference in Denver, STFC booth.

Community Activities

- Reviewer for ICLR 2025
- Co-reviewer for Nature Communications
- Organisational member for the Fundamentals of AI reading group at Queen Mary University of London
- Team captain and participate in the CIUK Cluster Challenge 2024

Publications

- [1] G. Mason-Williams and F. Dahlqvist, “What Makes a Good Prune? Optimal Unstructured Pruning for Maximal Cosine Similarity,” in *The Twelfth International Conference on Learning Representations*, 2024. [Online]. Available: <https://openreview.net/forum?id=jsvvPVVzwf>.
- [7] I. Mason-Williams, G. Mason-Williams, and M. Sandler, “Knowledge Distillation: The Functional Perspective,” in *NeurIPS 2024 Workshop SciForDL*, 2024. [Online]. Available: <https://openreview.net/forum?id=Cgo73ZnAQc>.

Pre-prints

- [5] G. Mason-Williams, D. Bond, and M. Basham, *GRAM - General RAM*, version 1.0.0, Sep. 2020. DOI: 10.5281/zenodo.4014691. [Online]. Available: <https://doi.org/10.5281/zenodo.4014691>.
- [6] G. Mason-Williams, D. Bond, and M. Basham, *DisTRaC - Distributed Transient Ram Ceph*, version 1.0.0, Sep. 2020. DOI: 10.5281/zenodo.4013776. [Online]. Available: <https://doi.org/10.5281/zenodo.4013776>.
- [8] G. Mason-Williams, D. Bond, and M. Basham, *DisTRaC: Accelerating High Performance Compute Processing for Temporary Data Storage*, 2022. DOI: 10.48550/ARXIV.2212.03054. [Online]. Available: <https://arxiv.org/abs/2212.03054>.

Talks

- [2] G. Mason-Williams, *An Introduction to DosNA: Distributed Numpy Arrays for High-performance cloud computing*, version 1, Mar. 2022. DOI: 10.5281/zenodo.6411813. [Online]. Available: <https://doi.org/10.5281/zenodo.6411813>.
- [3] G. Mason-Williams, “DisTRaC: Accelerating High Performance Data Processing,” English, Computing Insight UK 2022 : Sustainable HPC, CIUK ; Conference date: 01-12-2022 Through 02-12-2022, Dec. 2022. [Online]. Available: <https://www.scd.stfc.ac.uk/Pages/CIUK2022.aspx>.
- [4] G. Mason-Williams, “DisTRaC: Accelerating High-Performance Compute Processing for Temporary Data Storage,” English, Ceph Virtual 2022; Conference date: 03-11-2022 Through 16-11-2022, Nov. 2022. [Online]. Available: <https://ceph.io/en/community/events/2022/ceph-virtual/>.