BEN GUTTERIDGE

benpgutteridge@gmail.com ◇ bengutteridge.github.io

PROFILE

Machine learning researcher (DPhil, Oxford) working on learning from relational and structured data (graphs, molecules, sequences, documents), with a focus on long-range interactions, as well as long-context reasoning in LLMs. Industry experience in applied LLM research and observability, data science and software engineering.

EDUCATION

University of Oxford: DPhil in Machine Learning, *Keble College* **2021–2026**
Viva passed with minor corrections. EPSRC AIMS CDT (<5% acceptance). Thesis: *On the Range of Graph Neural Networks*; supervisors Prof. Xiaowen Dong and Prof. Michael Bronstein.

University of Oxford: MEng Engineering Science, *New College* **2017–2021**
First Class; Academic Scholar; dissertation on variational inference and Gaussian processes.

A*A*A*A at A-Level (2017), 10 A*s at GCSE (2015); state comprehensive

RESEARCH PROJECTS

Dynamically Rewired Message Passing Neural Networks, *ICML 2023, first author*
— Designed a graph neural network (GNN) architecture (layer-dependent rewiring with delayed skip connections between distant nodes) that captures long-range interactions without the quadratic cost and loss of graph inductive bias of full attention; matched or beat graph Transformers on four long-range benchmarks (molecular, peptide and segmentation graphs) and showed robust scaling with model depth, where standard GNNs break down.
— 170+ citations; featured in an ICML 2023 keynote and the ICML 2024 tutorial on graph learning.

Measuring Long-Range Interactions in GNNs, *ICML 2025, co-first author*
— Formalised the vague notions of range and long-range interactions, with quantitative range measures for models and tasks. Used them to audit popular benchmarks and showed that some are less long-range than assumed; the measures have since been adapted to sequence data for RL.

Architecture-Adaptive Graph Foundation Models, *GRaM at ICLR 2026, first author*
— Showed that current graph foundation models (GFMs) cannot generalise across task range, a limitation hidden by standard benchmarks, and built a model that *adapts its architecture* at inference time, zero-shot: better on average than existing GFMs, as well as GNNs trained *end-to-end* on each task, across long- and short-range tasks.

Multi-Modal LLMs for Handwritten Document Transcription, *COLM 2026, first author*
— First systematic comparison of zero-shot methods for multi-page handwriting transcription, on challenging documents (including archaic language and non-Latin script); released a new dataset. Proposed novel prompting methods: give an LLM the OCR of the whole multi-page document plus a *single page image*. This matched or beat methods using every page, at lower cost.

Detecting disease-carrying mosquitoes from phone recordings, *2020, 2021*
— Two summer research internships on the HumBug project with the Oxford ML Research Group.
— Built and evaluated a voice-activity-detection pipeline that removes speech from mosquito recordings (97% of speech removed, 75% of mosquito sound preserved).
— Contributed Bayesian CNN and ResNet models and experiments to the 2021 NeurIPS paper.

INDUSTRY EXPERIENCE

Deep Learning Intern, *QuantCo, London* **Jun–Aug 2024, returned Jul–Oct 2025**

— *2024, independent research project on LLM reasoning over long documents:* diagnosed the main failure modes in key entity extraction, compared prompting, few-shot and fine-tuning solutions, showing a small fine-tuned model matched the best large model at about 3% of the cost; presented to the deep learning team.

— *2025, LLM evaluation and MLOps:* implemented LLM tracing and evaluation with Langfuse for a large European health insurer's document-extraction pipeline (merged, in use) and for QuantCo's internal document processing product, then deployed shared instances for the company on Kubernetes/EKS, with a written guide and best practices, six demos and a company-wide talk.

Visiting Data Scientist, *BCG X, London* **Mar–May 2024**

— Worked as a software engineer on the British Airways Pathfinder flight-schedule optimiser, a tool used by BA's flight planners: shipped a "freeze" feature letting planners fix specific flights before running the optimiser (previously only an aircraft's whole schedule could be frozen), alongside many other PRs and code reviews in a shared production codebase.

Summer Intern, *QinetiQ, Malvern* **2018, returned 2019**

— Awarded student sponsorship, as well as two summer placements on a counter-drone radar team, analysing test data and developing MATLAB tools to optimise radar calibration.

PUBLICATIONS

First ➤ Can Graph Foundation Models Generalize Over Architecture?, *Workshop on Geometry-grounded Representation Learning and Generative Modeling, ICLR 2026*
First ➤ Judge a Book by Its Cover: Multi-Modal LLMs for Multi-Page Handwritten Document Transcription, *COLM 2026; Oral at AAAI 2025 workshop*
Co-first ➤ On Measuring Long-Range Interactions in Graph Neural Networks, *ICML 2025*
Third ➤ On Vanishing Gradients, Over-Smoothing, and Over-Squashing in GNNs, *NeurIPS 2025*
First ➤ DRew: Dynamically Rewired Message Passing with Delay, *ICML 2023*
Co-author ➤ HumBugDB: A Large-scale Acoustic Mosquito Dataset, *NeurIPS 2021 D&B, Oral*
Co-author ➤ HumBug: An Acoustic Mosquito Monitoring Tool for use on Budget Smartphones, *Methods in Ecology and Evolution 2021*

Invited talks: Learning on Graphs and Geometry reading group; G-Research ML Seminar Series.

OTHER EXPERIENCE AND SKILLS

— **Spring into Quant Finance**, G-Research (Apr 2023): selected for a competitive spring week programme in ML, data science and quantitative finance.

— **ML Project Leader**, Engineers Without Borders Oxford (2020–2021): ran a two-term ML programme for undergraduates; supervised two mini-projects (cellular-image analysis for drug discovery; EEG seizure prediction).

— **Reviewer** for several ML conferences and workshops.

Skills: Python for data science and machine learning (PyTorch, Hugging Face, NumPy/pandas); LLM fine-tuning and evaluation; MLOps; Git, testing and code review.

INTERESTS

Books, hiking, running, writing/game-mastering *D&D* campaigns, surveying pubs.