

MATHIS WEIL

London, UK | mathis.weil@outlook.com | mathisweil.github.io | linkedin.com/in/mathis-weil | github.com/mathisweil
Curriculum Vitae | Eligible for the UK Graduate visa, no sponsorship needed

RESEARCH INTERESTS

Machine learning under real-world constraints: compute-efficient large language model (LLM) inference, generative models for sequential decision-making, and federated learning.

EDUCATION

University College London, MSc Data Science & Machine Learning | Sep 2025 - Sep 2026

Thesis: Federated Gradient Boosting for Commercial Insurance Risk Modelling, with Ki (submitted September 2026; result pending). Module marks: Statistical Natural Language Processing (**80/100**), Open-Endedness & General Intelligence (**81/100**). Other modules: Applied Machine Learning, Information Retrieval & Data Mining, Bayesian Deep Learning, Introduction to Statistical Data Science, Digital Finance.

Queen Mary University of London, BSc Computer Science | Sep 2022 - Jun 2025

First Class Honours (83%); **Final-Year Prize** for Outstanding Academic Achievement. Final-year project in collaboration with IBM.

PUBLICATIONS

- "Return-Weighted ELBO Fine-Tuning Degrades Masked Diffusion Planners". Muhammad Ali Khan*, **Mathis Weil***, Ahmet H. Güzel, Jack Parker-Holder, Ilija Bogunovic. NeurIPS 2026 Workshop: Beyond Next-Token Prediction (BeNT), Sydney, December 2026. *Equal contribution.

WHITE PAPERS

- "The Human Transformation that will Power your Agentic Future". White paper, UCL School of Management Analytics Lab with Capgemini, May 2026 (co-author).

RESEARCH EXPERIENCE

Machine Learning Researcher, Ki Insurance | Jun 2026 - Sep 2026

Lloyd's first fully digital follow syndicate. MSc thesis research embedded in the underwriting team. Supervisor: Prof Philip Treleven (UCL).

- Built a federated XGBoost benchmark in which insurers share only summary statistics, never raw records; on public freMTPL data across ten simulated insurers it recovers a **median 0.99 of the pooled model's per-insurer gain**.
- Showed, under a pre-registered per-insurer protocol, that as rate multipliers push claim rates apart the shared model prices up to **7 of 10 insurers** worse than their own models, and that on this dial refitting one intercept per insurer removes the harm; proposed personalised federated boosting with per-insurer model selection.
- Worked alongside Ki's data science and actuarial teams on the non-IID heterogeneity of real insurer books, and wrote a threat model for federated trees in insurance.

RESEARCH PROJECTS

Train as You Evict: Eviction-Aware Fine-Tuning for Compressed KV Cache Inference | Jan 2026 - Apr 2026

UCL Statistical NLP (module graded 80/100). Team of 5: led the implementation and co-authored the report (TACL format; coursework). Code: github.com/mathisweil/evo-memory (built on the Sakana AI NAMM codebase).

- Showed that post-hoc KV cache eviction (75 to 85% pruning) induces train-test distribution shift, collapsing **F1 from 32.6 to 19.0** on long-context QA from LongBench (Llama-3.2-1B-Instruct).
- **Recovered F1 to 28.9** under the same pruning budget by proposing eviction-aware LoRA fine-tuning that exposes the model to eviction during training (PyTorch, Hugging Face, PEFT).
- Characterised how the model adapts, measured by JS divergence, hidden-state drift, attention-mass redistribution and Jaccard overlap of retained tokens.

Masked Diffusion Planners on Craftax and MiniHack (ReMDM) | Jan 2026 - Sep 2026

UCL Open-Endedness & General Intelligence (module graded 81/100), extended into the BeNT paper above. Coursework team of 4: lead developer of both codebases, sole presenter of the video walkthrough. Code: github.com/mathisweil/craftax-ReMDM-planner, [/minihack-ReMDM-planner](https://github.com/mathisweil/minihack-ReMDM-planner).

- Showed that return-weighted ELBO fine-tuning degrades a pretrained masked diffusion planner on Craftax Classic, **from 11.81 to 8.22**, with none of **25 conditions** (three seeds each) ending above the starting checkpoint (JAX).
- The return term is about half the imitation gradient, yet cutting it fivefold by clipping scores lower; most of the drop is a sampling mismatch: matching the evaluation settings shrinks it **from 3.68 to 0.26 points** on a same-host re-run.

- Built a remasking diffusion (ReMDM) planner for MiniHack that drafts 64-move plans in parallel and wins **48.5%** of in-distribution episodes (PyTorch).

SkillFindr: Semantic Search Engine for IBM SkillsBuild | Sep 2024 - May 2025

QMUL final-year dissertation with IBM (sole developer of the semantic search component). Supervisor: Aisha Abuelmatti.

Code and report: github.com/mathisweil/SkillFindr-SemanticSearchEngine.

- Achieved **Recall@20 = 0.899**, **MRR = 0.843** and **MAP@20 = 0.712** on a manually annotated 50-query ground-truth set with graded relevance judgements.
- Served queries at **27.6 ms** average latency versus 62.3 and 48.4 ms for the keyword and BM25 baselines, which returned results for only 4% and 6% of queries.
- Built the component end to end: Selenium + BeautifulSoup crawler for 1,200+ course records, hybrid relational-plus-vector store, retrieval API, and RAG integration with an IBM Granite LLM chatbot.

PROJECTS

Passage Retrieval and Learning-to-Rank at Scale | Jan 2026 - Apr 2026

UCL Information Retrieval & Data Mining coursework

- Built an inverted-index retrieval stack over **182,469 passages** implementing TF-IDF, BM25 and query-likelihood language models, plus a logistic-regression re-ranker combining fastText embeddings with lexical features.
- Edged tuned BM25 (**MAP 0.2063 vs 0.2040**) under **909:1** class imbalance by augmenting embeddings with the BM25 score, L2 regularisation and L2-normalisation.

TEACHING EXPERIENCE

Postgraduate Teaching Assistant, UCL | Sep 2025 - Jun 2026

- Led weekly practical sessions and assessed coursework for 150+ MSc students (Introductory Programming; Computer Architecture & Operating Systems; Software Engineering) and 130+ Year 2 BSc students (Compilers); marked the May 2026 Compilers examinations.

Computer Science Lab Demonstrator, QMUL | Sep 2024 - May 2025

- Supported 400+ undergraduates across Database Systems, Object-Oriented Programming, Operating Systems and Internet Protocols.
- **Demonstrator of the Year 2025** (one of 14 awarded across the Computer Science department); selected by the module lead, Dr Thomas Roelleke, to co-design new Database Systems coursework applying machine learning to information retrieval in databases.

AWARDS

- **Final-Year Prize** for Outstanding Academic Achievement, Queen Mary University of London, 2025.
- **Demonstrator of the Year 2025**, Queen Mary University of London (one of 14 across the Computer Science department).
- **1st place**, Bidwells Real Estate Planning Hackathon (UCL Team 1), against teams from eight universities including Cambridge, Nov 2025.
- **2nd of 20 teams**, UCL Business Games case competition sponsored by EY-Parthenon (financial modelling and strategy), 2025.

EXPERIENCE

Software Engineer, Groupe Prunay | Jul 2024 - Aug 2024

Louveciennes, France

- Built internal insurance account-management software and the group's investor and recruitment site.

SKILLS

Programming: Python (PyTorch, JAX, Hugging Face, PEFT, XGBoost, NumPy, pandas, scikit-learn), SQL (PostgreSQL, pgvector), R, Java, Git, Linux, Docker, GCP, Weights & Biases, pytest

Research: LLMs, diffusion models, reinforcement learning (RL), federated learning, gradient-boosted trees (GBDTs), LoRA fine-tuning, ablation and experimental design, natural language processing (NLP), information retrieval

Languages: French (native), English (fluent; Cambridge C1 Advanced, 184), German (intermediate)

REFERENCES

Available on request.