

# Tor Gunnar Høst Houeland

Welwyn Garden City, UK

+44 7823 521186

[houeland@gmail.com](mailto:houeland@gmail.com)

## EMPLOYMENT HISTORY

### DRW UK, working from home — (2023 - present)

Software Engineer

- **FICCO AI and Polaris (2024 - present)**
  - Working on Polaris, a firm-wide internal LLM-powered AI platform (with a Python backend hosted on Kubernetes and a web UI similar to claude.ai, but integrated with internal data and libraries)
  - Working on ML infrastructure including shared standards for MCP servers
  - Ran an LLM gateway for Cumberland/FICCO before we had a firm-wide solution
  - Member of committee identifying potentially disruptive AI applications, which explored and approved initial LLMs and AI-assisted coding tools for the firm
- **Cumberland/FICCO web UIs (2023 - 2024)**
  - Frontend lead for 2 other engineers, collaborating with traders to improve existing UIs and develop new information-dense and performant UIs
  - Developed a highly customizable risk system UI to visualize ground truth and model predictions for a variety of use cases, including Excel-style formulas and references

### Google UK, working from home — (2014 - 2023)

Senior Software Engineer

- **Driving Google Cloud reliability investments with data (2021 - 2023)**
  - Part of a cross-functional team collaborating across Google Cloud to collect, organize, present, and drive change in the organization based on reliability data
  - Developed an internal UI to identify important outage features through finer-grained expert ratings based on hypothetical outages
  - Collaborated with data analysts to develop a hierarchical Bayesian ML model for predicting outage severity in NumPyro/JAX, with a SHAP-based approach to visualize how input feature values contribute to predicted outage severity
  - Productionized the model and integrated with other systems to predict severity and feature importance explanations for all outages across Google Cloud, and developed the UI to display it
- **Machine Learning Crash Course with TensorFlow APIs (2016, 2018, 2021)**
  - Instructor within Google for what is now a public [machine learning crash course](#)
- **Public status dashboard system (2019 - 2021)**
  - Highly-available status dashboard system, used across Google to notify the public of outages (e.g. for [Cloud](#), [Ads](#), [Search](#), [Workspace](#))
  - Collaborated across Google product areas to prioritize overall feature roadmap, rescope features to meet deadlines, and launch dashboards they needed on time
  - Designed the overall architecture, API, internal UI, and localization approach
  - Led a team of 5 other engineers across 2 sites to implement the solution
- **Incident response and management system (2018 - 2020)**
  - Frontend tech lead, ramped up UI engineering team from 2 engineers in London to 14 engineers across 3 sites

- Defined frontend code architecture and best practices, and implemented the majority of the initial web UI when the project first started
- Worked with product managers and UX interaction designers to define new features and refine the UI to match the needs of production engineers
- **Compute Engine SRE (2017)**
  - 6-month rotation to learn about Google Production and SRE
  - On-call for [Google Cloud Compute Engine](#), Google's external VM offering that runs on Borg (the predecessor to Kubernetes), and worked on internal monitoring
- **Knowledge management system for Ads Sales (2014 - 2017)**
  - Rewrote search functionality within an internal content management system: defined a search quality metric, used Google search tooling to develop a better-scoring ranking function, and translated it into features that we could efficiently support within our architecture
  - Increased search quality score by 40%, reduced search latency by 85%, and increased indexing speed by 3000%. Search improvements were pointed out as a major contributor to customer satisfaction increasing from 32% to 85% during this period
  - Contributed to a recommendation feature to suggest relevant documents to users, migrating from a handwritten scoring function to training a two-tower recommendation model

## **Department of Computer and Information Science at NTNU, Trondheim, Norway (2013)**

- Assignment coordinator for "Machine Learning and Case-Based Reasoning" course

## **INDEPENDENT ML RESEARCH**

### **Google TPU Research Cloud (2022-2025)**

- Designed a purpose-built synthetic dataset and structured probe evaluations to track emergent learning stages and training instabilities during transformer training
- Ported Llama 3.1 from PyTorch to JAX; trained a byte-level transformer LLM from scratch on TPUs; began multi-host port using a custom Ray-based distributed coordination library
- Ran 75k+ CIFAR-10 training experiments to systematically characterise optimizer and hyperparameter dynamics across training runs (approx. 20GB of results data)
- Published results and interactive visualizations at <https://houeland.github.io/>

## **EDUCATION**

### **Norwegian University of Science and Technology, Trondheim, Norway (2020) — PhD degree in Artificial Intelligence**

Doctoral thesis title: [Automated lazy metalearning in introspective reasoning systems](#)

Keywords: lazy learning, meta-level learning, ensembling, reinforcement learning

Publications: 11 papers, 72 citations: <https://scholar.google.com/citations?user=NwczZqUAAAAI>

### **Norwegian University of Science and Technology, Trondheim, Norway (2008) — Master's degree in Computer Science**

Master thesis title: [Reuse of Past Games for Move Generation in Computer Go](#)

Keywords: machine learning, case-based reasoning, UCT algorithm, Go board game