

Collin Aldaibis

AI ENGINEER

AI engineer with an MSc in Computing Science. I build production AI systems end to end, from training and fine-tuning language models to the web apps and infrastructure that ship them, and enjoy working directly with stakeholders to turn business goals into working software.

Utrecht, Netherlands / caldaibis@gmail.com / GitHub / LinkedIn

■ EXPERIENCE

○ Machine Learning Engineer at Netherlands Forensic Institute (NFI) Oct 2024 - Present

The Hague, Netherlands

Building forensic AI within a 40+ person data science guild, leading the engineering and LLM work across Scrum projects with the police, the Hansken team, and the biological and chemical data labs.

- Fine-tuned Whisper for Dutch forensic audio (ethnolects, dialects, and low-quality recordings), cutting word error rate from ~40% to ~15% on private forensic datasets and teaching it previously unseen slang.
- Pre-trained a Dutch ModernBERT from scratch on Dutch FineWeb / FineWeb-Edu to close a gap in open-source Dutch encoder models with a modern BERT/Roberta successor.
- Built and evaluated the Hansken Copilot RAG application, helping investigators search cases of millions of digital traces, fully on-prem with local embedding models and LLMs (case data may never leave the NFI), evaluated with RAGAS on hand-curated forensic datasets.
- Benchmarked Dutch LLMs (GPT-NL, Mistral, Gemma, Qwen) with SURF & TNO; contributed the first Dutch bias benchmark to EuroEval and built a fine-tuning/eval pipeline for a forensic text-classification use case to assess whether GPT-NL could outperform existing models.
- Established local LLM infrastructure for running and fine-tuning models on sensitive data (llama.cpp, Unsloth, Hugging Face, PyTorch), including sandboxed local coding agents and optimised training runs on NVIDIA H200s.
- Founded and chaired the Python Dev Club and LLM Club to raise engineering and LLM literacy across the guild (modern tooling like Ruff and uv/pdm, plus safe local agents), and advised NFI leadership on rolling out LLM technology organisation-wide.

Python · PyTorch · Hugging Face · Unsloth · llama.cpp · RAG · RAGAS · NVIDIA H200

○ Full-Stack Software Engineer at Cordys Analytics 2023 - 2024

Utrecht, Netherlands

Built the company's core AI product, giving clinicians and patients seamless access to AI-powered ECG diagnostics.

- Developed the diagnostics platform end to end, from model serving to clinician-facing UI.
- Supervised several international AI master's students on deep learning for ECG research theses.
- Organised company events and team activities.

C# · Python · PyTorch Lightning · React (TypeScript) · Docker · Bazel · Terraform · Azure · MLOps

○ Software & Machine Learning Engineer at Schiphol (via Incentro) 2022 - 2023

Amsterdam, Netherlands

Delivered ML and full-stack applications for Schiphol as a consultant, owning the relationship with the customer.

- Built and maintained a Flask ML service, a Laravel backend, and an Angular frontend.
- Worked directly with stakeholders at Schiphol to translate business objectives into delivered software.
- Applied NLP techniques, word embeddings, and LLMs to production use cases.

Python · Flask · Laravel (PHP) · Angular (TypeScript) · GCP · Docker · MySQL · CI/CD

■ PERSONAL PROJECTS

Aparecium: Agentic LLM Benchmark Extraction ↗

A multi-stage agentic pipeline that extracts AI model benchmark scores from official model pages, model cards, and benchmark charts: web search, content/table parsing, OCR on chart images, LLM-based extraction, and resolution against a canonical benchmark catalog, evaluated against ground truth. Includes a self-improving agent that diagnoses recurring failure modes, applies targeted fixes, and re-validates, plus a frontend for exploring and comparing results.

Python · LLM agents · OCR · TypeScript · uv

Speechless: Whisper Vocabulary Adaptation ↗

Research pipeline adapting Whisper to out-of-vocabulary terms (e.g. Dutch medical vocabulary) without training on real audio: a single-codebook residual VQ over frozen Whisper encoder states and a GPT-2-style decoder mapping text to RVQ token sequences. Reproduces and extends the Speechless (2025) methodology.

Python · PyTorch · Whisper · Residual VQ · ASR

Autonomous Browser Agent ↗

End-to-end agentic automation: triggered by incoming email, it opens a target listing, extracts metadata, and hands off to a browser agent that navigates third-party sites, writes a tailored message from a template, fills the application form, uploads documents, and submits, fully autonomously.

Python · LLM agents · Browser automation

■ EDUCATION

Utrecht University

2018 – 2022

MSc, Computing Science

Master's thesis "Investigating relaxed probability updating games" (grade 8.2). Coursework: advanced machine learning, data mining, pattern recognition, evolutionary computing, multi-agent learning, network science.

Utrecht University

2015 – 2018

BSc, Computing Science

Coursework: algorithmics, optimisation, concurrency, security, graphics, computational intelligence, functional programming.

■ CERTIFICATIONS

AWS Certified Machine Learning (Specialty) · Amazon Web Services

2023

Associate Cloud Engineer · Google Cloud

2023

Deep Learning Specialisation · DeepLearning.AI (Coursera)

2022