

DANIEL DAVIES

ML Research Engineer – Pre-Training & Interpretability · Full-Stack Systems
danieljdavies.com · github.com/DanielJamesDavies · danieljamesdavies12@gmail.com
+44 7823 777530 · London, UK

PROFILE

ML Research Engineer who runs careful experiments, learns fast, and ships production AI. MSc research in LLM pre-training, synthetic data generation, developed novel mechanistic interpretability algorithms and tooling designed for efficiency. Comfortable leading end-to-end from hypothesis to evals and iteration.

Highlights: Trained 254M-param LLM in <2 GPU days (val. loss ~1.8). Generated 2B-token synthetic dataset. Built latent explorer (turingexplorer.com) with optimised TTFB <1s, search 491,520 latents in ~3s (≤9s), and circuit generation ~3s. MSc Artificial Intelligence with Distinction. Shipped 40+ production apps/automations, common flows cut from hours/days to minutes/seconds.

SKILLS

- **ML/AI:** LLM pre-training & distillation, synthetic dataset generation, mechanistic interpretability (SAEs, feature labelling, circuit discovery), evaluation, prompt/agent systems, RAG, structured decoding, function calling, exploratory data analysis.
- **Languages/Frameworks:** Python (PyTorch, TensorFlow, NumPy, FastAPI), JavaScript/TypeScript (React, Node), SQL, HTML/CSS, R, HF Transformers.
- **Infra & Observability:** GCP, Azure, Heroku, Linux, Git, CI/CD, metrics/logs/traces, cost control.

EXPERIENCE

Applied-AI Full-Stack Software Engineer (Independent Contractor) with Projekt Rising Ltd

Sep 2022 – Present (3 yrs)

- Architected and shipped **40+** production AI applications and automations for SMEs, streamlining key workflows from **hours/days to minutes/seconds**. Many regular users per application.
- Collaborated across a growing team, coordinated projects, made key technical decisions (coding standards, languages, frameworks), and improved delivery cadence while keeping incident rate and cost low.
- Content generation, deep search systems, document pipelines, recommendation, dashboards, RAG/agent services. React, TypeScript, Python, FastAPI, GCP/Azure, SQL/vector stores, CI/CD, cost monitoring.

EDUCATION

MSc Artificial Intelligence (Distinction)

Brunel University of London | September 2023 – December 2024

BSc Computer Science (Artificial Intelligence, First-Class Honours)

Brunel University London | September 2019 – July 2022

BTEC Level 3 Extended Diploma in IT & GCSEs

West Herts College | September 2017 – June 2019

Triple Distinction Star (D*D*D*)

SELECTED PROJECTS

Turing-LLM Explorer

turingexplorer.com | github.com/DanielJamesDavies/Turing-LLM-Explorer

Interactive mechanistic interpretability tool for Turing-LLM-1.0-254M. Navigate sparse autoencoder latent space with ~3s (≤ 9 s) search. Run completions with time to first token < 1 s and visualise latent circuits generated in ~3s. Key Languages/Frameworks: Python, PyTorch, NumPy, Flask, React, JavaScript, and more.

On Novel Approaches Towards Interpretability: Training and Understanding Turing-LLM-1.0-254M

github.com/DanielJamesDavies/Turing-LLM-1.0-254M

Trained Turing-LLM-1.0-254M from initialisation (val. loss ~ 1.8). Experimented with LLM architectures. Implemented synthetic dataset generation system, with its output of 2 billion tokens used to train the LLM. Developed and evaluated novel mechanistic interpretability algorithms for feature labelling and circuit discovery. Trained 12 TopK sparse autoencoders (for each layer) on LLM MLP layer activations for feature extraction. Key Languages/Frameworks: Python, PyTorch, NumPy, Hugging Face Transformers.

Turios (Humanoid Robot)

Building a low-cost full-sized humanoid robot as a scalable real-world data collection strategy to be used in model training, finetuning, and interpretability research. Designed high-torque actuators with FOC motor control, involving creating cycloidal drives, soldering encoders, and writing code for position control loops. Built a tendon-driven hand and forearm capable of handling heavy load.

Atlas Story App

atlas-story.app | github.com/DanielJamesDavies/AtlasStoryApp

Developed an extensive full-stack platform for story creation and sharing. Includes fully customizable interactive 3D and 2D maps, character relationship charts, Spotify integration, structured story content, auth. Key Languages/Frameworks: React, Node.js, Three.js, MongoDB, Heroku.

Algonet / The Prediction of Energy Consumption of Algorithm Implementations

github.com/DanielJamesDavies/Algonet

BSc dissertation. Trained a model to predict energy consumptions of algorithms. Data collected by running sets of algorithms (e.g. search, sort, graphs) with different implementations (e.g. bubble sort, binary search) on devices and recording their energy consumption. Key Languages/Frameworks: Python, PyTorch, NumPy, scikit-learn, SciPy, JavaScript.

REFERENCES

References available on request.