Joshua NANCEY

Machine Learning, Computer Vision & LLM Engineer

M.Eng. UTC (**Mod'Math Honors**) • Deep Learning, Grounded LLMs & MLOps

joshua.nancey@etu.utc.fr • +33 7 49 15 69 60
github.com/joshdeutc • joshdeutc.github.io
linkedin.com/in/joshua-nancey
Paris, France (Open to Relocation: EU / CH / UK)



Professional & Research Experience

▸ **Epiconcept — R&D Data Science Lab** *March 2026 – Sept. 2026 (6 months)*

Deep Learning, Medical Vision & LLM Research Intern *Paris, France*

- **Grounded Clinical LLM Extraction (ECDC / MIMIC-IV):** Deployed and benchmarked 20 local LLM configurations (**MedGemma-4B/27B**, **Gemma-2-27B**, **Qwen-2.5-32B**, **Mistral-Small-24B**) quantized in 8/4-bit under a single **NVIDIA A100 40 GB** budget (**F1 = 0.91**).
- **Anti-Hallucination & Cross-Lingual Transfer:** Built a deterministic substring verifier (`evidence_is_grounded`) enforcing verbatim quotes and strict JSON schemas; achieved **F1 = 0.88** in zero-shot cross-lingual transfer to Croatian clinical notes via **NLLB-200** / **EuroLLM-9B**.
- **High-Resolution Medical Vision (GMIC):** Fine-tuned the interpretable **GMIC** network (**ResNet-22** global saliency + **ResNet-18** local patch attention) in 16-bit precision on 300+ GB of RSNA mammograms (2944×1920), boosting AUC from 0.683 to **0.889** (95% CI: [0.858, 0.921]).
- **Statistical Risk Control:** Implemented selective prediction (*SGR* framework, Hoeffding bounds, **MC Dropout**, **Deep Ensembles**) guaranteeing FNR ≤ 5% and FPR ≤ 5%; co-authored an 8-chapter interactive **Docker** / **Jupyter** training capsule.

▸ **BMBI Laboratory (CNRS UMR 7338 / UTC)** *Sept. 2025 – Jan. 2026 (5 months)*

Graduate Research Project (TZ2) — 3D/4D Implicit Neural Surfaces (SIREN) & Autograd Geometry *Compiègne, France*

- **Space-Time Implicit Surfaces $f_\theta(x,y,z,t) = 0$:** Modeled 4D deforming biological capsules using **SIREN** networks, Fourier features, and **Eikonal regularization** ($||\nabla f|| = 1$), achieving **99.65% IoU**.
- **Analytical Parsimony & Differential Geometry:** Recovered the analytical Kuchel quartic in **10 coefficients** (vs. 197,886 neural weights) and verified the **Gauss-Bonnet theorem** to **0.025%** via 2nd-order PyTorch `autograd`.

▸ **Aubay — Innovation Department** *Sept. 2024 – Feb. 2025 (6 months)*

Data Science & AI Intern — Contextual Anomaly Detection *Boulogne-Billancourt, France*

- **Deep Tabular Learning:** Benchmarked **GANDALF**, **FT-Transformer**, and **TabTransformer** (PyTorch Tabular) coupled with **Isolation Forest** / **DIF** on mixed datasets (**F1 = 0.903**, **ROC-AUC = 0.985**).
- **Full-Stack Engineering:** Delivered a production-ready web platform with a **Flask REST API**, **React** frontend, **AWS S3** storage, and reproducible **Conda** / **Docker** containers.

Education

▸ **Master of Engineering (M.Eng.) in Computer Science, AI & Mathematical Modeling** *2023 – 2026*

Université de Technologie de Compiègne (UTC) *Compiègne, France*

- **Data & AI Track:** Deep Learning, Computer Vision (*SY32*), Statistical Learning & Clustering (*SY09*), ETL & Data Warehousing (*NF26*).
- **Mod'Math Honors:** Convex & Combinatorial Optimization (*RO04*), Stochastic Processes & Markov Chains (*RO05*), Numerical Analysis (*MT09*).

▸ **Preparatory Classes for Grandes Écoles (CPGE PTSI / PT — Intensive Math & Physics)** *2020 – 2023*

Lycée Raspail *Paris, France*

Selected AI & Systems Engineering Projects

▸ **Jarvis & Simulation-Go (IA04) — Grounded Multi-Judge RAG Engine & Multi-Agent System** *Python, FastAPI, SQLite FTS5*

- Local **SQLite FTS5 BM25** RAG knowledge base, 51-board ATS harvester, and 3-LLM grounded evaluation jury; concurrent autonomous multi-agent simulation engine in **Go (Golang)**.

▸ **MonLycée Cloud Auth Relay & Custos — Playwright Session Broker & Android Security** *FastAPI, Playwright, Kotlin*

- Cloud-deported Keycloak SSO session broker (< 50 ms RAM cache, **Edge Manifest V3** extension) and native **Android Device Owner** security suite in Kotlin.

▸ **SY09 & SY32 — Statistical Learning, Wald Testing & Classical Computer Vision (HOG/SVM)** *Python, Scikit-Learn, OpenCV*

- Logistic regression with **Wald statistical tests**, **Gower** / **Adaptive Mahalanobis K-Means** on cardiac datasets, and sliding-window object detection via HOG features, SVM/Random Forest, hard negative mining, and NMS.

Technical Skills

AI Frameworks : PyTorch, Hugging Face (Transformers, bitsandbytes), PyTorch Tabular, Scikit-Learn, OpenCV, SIREN

MLOps & Systems : Docker, Linux (systemd), Git/CI, FastAPI, Flask, AWS S3, SQLite FTS5

Languages : Python, C/C++17 (Qt 6), SQL, Kotlin, R, Scilab, Bash, JavaScript (React / Manifest V3), LaTeX / Quarto

Languages & Interests : French (Native), English (Fluent B2+), Boxing, Freestyle Wrestling