Joshua NANCEY

Ingénieur R&D — Machine Learning, Vision & LLMs

Ingénieur UTC (**Label Mod'Math**) • Deep Learning, LLMs & Déploiement

joshua.nancey@etu.utc.fr • +33 7 49 15 69 60
github.com/joshdeutc • joshdeutc.github.io
linkedin.com/in/joshua-nancey
Paris / Europe (Mobile) • Permis B



Expériences Professionnelles & Recherche R&D

‣ Epiconcept — Équipe R&D Data Science *Mars 2026 – Sept. 2026 (6 mois)*

Ingénieur R&D en Apprentissage Profond, Vision Médicale & LLMs *Paris, France*

- **Extraction clinique par LLMs locaux (Projet ECDC / MIMIC-IV) :** Déploiement et benchmark de 20 configurations de LLMs (**MedGemma-4B/27B**, **Gemma-2-27B**, **Qwen-2.5-32B**, **Mistral-Small-24B**) quantifiés en 8/4 bits sous budget strict NVIDIA A100 40 Go (**F1 = 0,91**).
- **Vérification anti-hallucination & transfert cross-lingue :** Protocole imposant des citations verbatim (`evidence_is_grounded`) et JSON strict ; transfert zéro-shot vers des notes cliniques croates via **NLLB-200** / **EuroLLM-9B** (**F1 = 0,88**).
- **Vision par ordinateur haute résolution (GMIC) :** Fine-tuning du réseau global/local à attention **GMIC** (**ResNet-22** + **ResNet-18**) en 16 bits sur 300+ Go de mammographies RSNA (2944×1920), portant l'AUC de 0,683 à **0,889** (IC 95% : [0,858 ; 0,921]).
- **Contrôle statistique du risque :** Classification sélective (cadre SGR, bornes de Hoeffding, **MC Dropout**, **Deep Ensembles**) garantissant FNR ≤ 5% et FPR ≤ 5% ; capsule interactive de 8 chapitres sous **Docker**.

‣ Laboratoire BMBI (CNRS UMR 7338 / UTC) *Sept. 2025 – Janv. 2026 (5 mois)*

Projet de Recherche (TZ2) — Réseaux Implicites 3D/4D (SIREN) & Géométrie Différentielle *Compiègne, France*

- **Surfaces implicites spatio-temporelles $f_\theta(x,y,z,t) = 0$:** Reconstruction 4D de membranes déformables par réseaux **SIREN**, encodage de Fourier et régularisation **Eikonale** ($||\nabla f|| = 1$), atteignant **99,65% d'IoU**.
- **Géométrie différentielle & parcimonie :** Retrouve l'équation analytique de Kuchel en **10 coefficients** (vs 197 886 poids neuronaux) et calcule les courbures de Gauss/moyenne par `autograd` d'ordre 2 (théorème de Gauss-Bonnet vérifié à **0,025%**).

‣ Aubay — Pôle Innovation *Sept. 2024 – Févr. 2025 (6 mois)*

Ingénieur Data Science & IA — Détection d'Anomalies Contextuelles *Boulogne-Billancourt, France*

- **Deep Learning Tabulaire :** Benchmark de **GANDALF**, **FT-Transformer** et **TabTransformer** (PyTorch Tabular) couplés à **Isolation Forest** / **DIF** sur données mixtes (**F1 = 0,903**, **ROC-AUC = 0,985**).
- **Industrialisation Full-Stack :** Développement d'une API REST **Flask**, interface **React**, connecteur **AWS S3** et environnement reproductible **Conda** / **Docker**.

Formation Académique

‣ Diplôme d'Ingénieur en Génie Informatique — Filière IAD & Label Mod'Math *2023 – 2026*

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

- **Data & IA :** Deep Learning, Vision par Ordinateur (*SY32*), Apprentissage Statistique & Clustering (*SY09*), Data Warehouse & ETL (*NF26*).
- **Label Mod'Math :** Optimisation convexe & combinatoire (*RO04*), Processus Stochastiques & Markov (*RO05*), Analyse Numérique (*MT09*).

‣ Classes Préparatoires aux Grandes Écoles (CPGE PTSI / PT) *2020 – 2023*

Lycée Raspail *Paris, France*

Projets Personnels & Systèmes IA

‣ Jarvis & Simulation-Go (IA04) — Moteur RAG Multi-Juges (FTS5) & Systèmes Multi-Agents *Python, FastAPI, SQLite FTS5*

- Moteur RAG local (**SQLite FTS5 BM25** + graphe d'entités), agrégateur de 51 APIs ATS et comité de 3 LLMs à preuves verbatim ; simulateur de systèmes multi-agents autonomes concurrents en **Go (Golang)**.

‣ MonLycée Cloud Auth Relay & Custos — Broker SSO Playwright & Sécurité Android *FastAPI, Playwright, Kotlin*

- Relais Cloud d'authentification Keycloak (cache mémoire < 50 ms, extension **Edge Manifest V3**) et suite Android **Device Owner** en Kotlin avec watchdogs anti-contournement.

‣ SY09 & SY32 — Apprentissage Statistique, Tests de Wald & Vision Classique (HOG/SVM) *Python, Scikit-Learn, OpenCV*

- Régression logistique avec tests de **Wald**, K-Means à distance de **Gower** / **Mahalanobis** sur données cardiaques, et détection d'objets par descripteurs HOG, SVM/Random Forest, *hard negative mining* et NMS.

Compétences Techniques

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

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

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

Langues & Loisirs : Français (natif), Anglais (courant B2+), Boxe anglaise, Lutte libre