Joshua NANCEY

Ingénieur Modélisation Mathématique, RO & R&D IA

Ingénieur UTC (**Label Mod'Math**) • Optimisation, Processus Stochastiques & Garantie de Risque

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



Formation Académique

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

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

- **Label Modélisation Mathématique (Mod'Math) :** Optimisation continue et discrète (Simplexe, Dualité, conditions KKT, Lagrangien augmenté, descente de gradient — *RO04*), Recherche Opérationnelle & Modélisation Stochastique (Chaînes de Markov, processus de Poisson, files d'attente, simulation d'événements discrets — *RO05*), Analyse numérique matricielle & EDO (SVD, QR, Runge-Kutta — *MT09*).
- **Statistiques & Intelligence Artificielle :** Estimation paramétrique, tests d'hypothèses, intervalles de confiance (*SY02*), Apprentissage statistique & bayésien, Analyse de données multidimensionnelles (*SY09*), Vision par ordinateur (*SY32*).

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

Lycée Raspail *Paris, France*

Expériences Professionnelles & Recherche R&D

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

Ingénieur Recherche — Garantie Statistique du Risque, Vision Médicale & LLMs *Paris, France*

- **Classification sélective à risque garanti (cadre SGR & bornes de Hoeffding) :** Implémentation et extension bi-seuil (τ_-, τ_+) des algorithmes de Geifman & El-Yaniv pour borner simultanément le taux de faux négatifs (FNR ≤ 5%) et de faux positifs (FPR ≤ 5%) en échantillon fini avec garantie probabiliste 1–δ ; quantification d'incertitude par **MC Dropout** et **Deep Ensembles**.
- **Apprentissage profond 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, prévalence 2,1%), portant l'AUC de 0,683 à **0,889** (IC 95% : [0,858 ; 0,921]).
- **Inférence LLM contrainte & vérification de preuves (Projet ECDC) :** Benchmark de 20 configurations de LLMs (**MedGemma-4B**/**27B**, **Gemma-2-27B**, **Qwen-2.5-32B**) sous budget GPU NVIDIA A100 40 Go sur notes cliniques MIMIC-IV ; vérificateur déterministe de citations verbatim (`evidence_is_grounded`, **F1 = 0,91**) et transfert cross-lingue zéro-shot vers le croate (**F1 = 0,88**).
- **Transfert méthodologique :** Co-auteur d'une capsule interactive de 8 chapitres (Quarto / Jupyter / Docker) sur la maîtrise du risque en IA médicale.

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

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

- **Représentations implicites spatio-temporelles $f_\theta(x,y,z,t) = 0$:** Modélisation continue de membranes biologiques déformables (globules rouges de Kuchel et micro-capsules sous cisaillement) par réseaux **SIREN** ($\sin(\omega_0 \cdot)$), encodage de Fourier et régularisation **Eikonale** ($||\nabla f|| = 1$), atteignant une **IoU de 99,65%** sur déformations non-affines.
- **Parcimonie mathématique vs capacité neuronale :** Démonstration que la régression polynomiale de degré 4 retrouve l'équation exacte de Kuchel en **10 coefficients** (résidu nul aux erreurs machine près) contre 197 886 poids pour le MLP.
- **Géométrie différentielle analytique :** Calcul exact des courbures de Gauss (K) et moyenne (H) par différentiation automatique PyTorch d'ordre 2 (`autograd`) sans maillage ; validation numérique du théorème de **Gauss-Bonnet** ($\oint K\, dA = 4\pi$) à **0,025%** près.

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

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

- **Modélisation d'anomalies sur données tabulaires mixtes :** Étude comparative de modèles profonds (**GANDALF**, **FT-Transformer**, **TabTransformer** via PyTorch Tabular) et d'arbres d'isolation (**Isolation Forest**, **DIF**) pour séparer variables contextuelles et comportementales (**F1 = 0,903**, **ROC-AUC = 0,985**).
- **Architecture logicielle :** Développement d'une plateforme d'audit full-stack (API REST **Flask**, **React**, stockage **AWS S3**, **Docker**).

Projets Scientifiques & Ingénierie Système

▸ **SY09 — Apprentissage Statistique, Tests de Wald & Clustering Adaptatif** *Python, Scikit-Learn, Statsmodels*

- Régression logistique avec tests statistiques de **Wald**, classifieurs bayésiens/euclidiens et clustering **K-Means à distance de Gower** (données mixtes) et métrique adaptative de **Mahalanobis** sur données cliniques cardiaques.

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

- Moteur d'évaluation multi-juges LLMs avec vérification déterministe d'extraits verbatim (**SQLite FTS5 BM25**) et simulateur de systèmes multi-agents autonomes concurrents en **Go (Golang)** avec visualisation temps réel.

Compétences Techniques

Mathématiques & RO : Optimisation convexe/combinatoire, Markov, Poisson, Files d'attente, Tests statistiques, Bornes de Hoeffding, Autograd d'ordre 2

Deep Learning & IA : PyTorch, Hugging Face, PyTorch Tabular, Scikit-Learn, SIREN/PINNs, OpenCV

Langages : Python, C/C++17 (Qt 6), SQL, R, Scilab/MATLAB, Kotlin, Bash, LaTeX / Quarto

Systèmes & Langues : Linux, Docker, Git, FastAPI/Flask, Français (natif), Anglais (professionnel B2+)