

HICHEM GOUIA

AI ENGINEER (AGENTIC, GENAI, NLP, COMPUTER VISION)

Clichy 92110, France | contact.hichemgouia@gmail.com | 0670718977 | [Linkedin](#)



PROFILE

AI Engineer with hands-on experience delivering production-grade ML/AI systems and scalable data pipelines. Skilled in MLOps, Agentic, NLP, Computer Vision, data engineering, Python, cloud platforms and full-stack development; Master's in Computer Science (Applied AI & Data) and MasterDevFrance 2025 finalist.

AWARDS & HONORS

MasterDevFrance 2025 Finalist (AI/Algorithm Hackatons)

PROFESSIONAL EXPERIENCE

Software Engineer, AI/Data, EstimerMonCommerce.fr Jan 2024 to Sept 2026
Full-Remote

- Implemented a production-grade OCR pipeline to extract accounting data and autopre-fill the estimation workflow, ensuring high reliability and availability; instrumented model observability and performance metrics using LangFuse.
- Led R&D and prototyping of generative AI solutions (LLMs, vLLMs), conducted comparative analysis of open-source LLMs (Mistral, Deepseek) to optimize performance/cost ratio, and orchestration tools (LangChain) to accelerate product innovation.
- Designed and developed an automated web-scraping ingestion pipeline for real-estate listings (raw ingestion, cleaning, normalization, scalable storage), producing consistent datasets for analytics and modeling.
- Created an interactive geospatial market-analysis module (data pipeline, normalization, visualization) to support local market studies and product decision-making.
- Contributed to full-stack development and deployment of SaaS features using React (frontend) and PHP (backend), improving delivery cadence and product reliability.

App Developer, LTG Services Mar 2023 to Jun 2023

- Designed, built and shipped a cross-platform iOS/Android professional networking mobile app, owning architecture, UX and end-to-end production deployment to app stores; implemented onboarding flows and analytics to measure user engagement.

EDUCATION

Epitech Technology Sept 2024 to Sept 2026
Master in Computer Science - Applied AI & Data, Nancy

- Master's program emphasizing end-to-end development and deployment of AI systems for real-world applications.
- Selected coursework: Machine Learning; Deep Learning; Reinforcement Learning (Deep QLearning); Natural Language Processing (NER, Transformers/BERT); Speech & Audio Processing.
- Practical focus: transfer learning, model fine-tuning, data preprocessing & augmentation, evaluation metrics, model monitoring and deployment pipelines.
- Capstone and lab projects delivering production-oriented models with attention to scalability, robustness, and evaluation.

Epitech Technolgy Sept 2022 to Aug 2024
Bachelor IT System, Nancy

- Bachelor program emphasizing systems engineering, networking, and practical IT infrastructure.
- Core topics: Linux administration, networking, scripting (Python/Bash), databases, and cloud fundamentals.
- Practical projects: automated deployment, system monitoring, and small-scale service orchestration.

AREAS OF EXPERTISE

Scientific Computing & AI Neural Networks	Machine learning MLOps	Data Engineering Cloud	Programming Project Management
--	---------------------------	---------------------------	-----------------------------------

PROJECTS (GITHUB.COM/IAMHMH)

Automated Skin Lesion Classification (Computer Vision Research) : Conducted a comparative analysis of CNN architectures (EfficientNet vs. ResNet) using Transfer Learning to classify skin pathologies from clinical imagery. Engineered a robust preprocessing pipeline implementing stain normalization and advanced data augmentation techniques to mitigate dataset imbalance and overfitting.

Advanced NER & Language Modeling (NLP Research) : Explored and implemented state-of-the-art NLP architectures, specifically focusing on Transformers and BERT-based models for Named Entity Recognition (NER) tasks. Conducted experiments in transfer learning and model fine-tuning to adapt pre-trained language models to specific domain datasets. Evaluated system performance using rigorous NLP metrics, ensuring a deep understanding of attention mechanisms and sequence labeling challenges.

Audio Stem Extraction Project : Designed an automated ingestion pipeline featuring spectral analysis, audio normalization, and waveform preprocessing to optimize inputs for neural inference. Architected a scalable, GPU-accelerated inference system using serverless orchestration to handle high-dimensional audio data with low latency.