

Adrien SKRZYPCZAK

Engineering student and Data/AI Analyst Intern at AI Builders (Paris).
Pursuing a Master's in AI & Data at the intersection of NLP, Machine Learning, and Generative AI.

📍 Paris & Hauts-de-France, France - ✉ adrien.skrzyp@gmail.com - 🔗 adrien-skr.dev - 🌐 adrienskrzypczak
🔗 AdrienSkr

PROFESSIONAL EXPERIENCE

Data/AI Analyst Intern, AI Builders – Paris, France Sept 2026 – Feb 2027

- Monitor LLMs and agentic AI trends, compare Data/AI solutions, and develop proof-of-concept prototypes.

Generative AI Intern, Nicolaudie Group – Montpellier, France June 2025 – Aug 2025

- Design a natural language to JSON configuration transformation system by combining entity extraction, RAG, and reranking with FAISS, LangChain, and OpenAI.
- Industrialize a REST API and prototype a Streamlit interface, containerized with Docker.
- Build and annotate training and test datasets with Label Studio, then generate initial predictions to accelerate the annotation process.

EDUCATION

University of Sherbrooke, Master's in AI and Data Science – Sherbrooke, Canada Aug 2025 – Aug 2026

- Machine Learning and methodology in Python: supervised and unsupervised learning, preprocessing, cross-validation, and model optimization.
- Deep Learning, NLP, and advanced AI: neural networks (CNN, RNN, LSTM, GRU), Transformers, language models, reinforcement learning, recommendation.

Engineering School of the Opal Coast, Engineering Degree in Computer Science – Calais Sept 2023 – Sept 2026

- Solid foundation in signal processing, image processing, AI, and applied statistics.
- Management of group projects in academic and associative contexts, using agile methodology.

PROJECTS

SemEval 2025 Task 9 – Explainable Food Hazard Detection [🔗](#) Jan 2026 – Apr 2026

- Design a multi-task NLP model based on RoBERTa-large to classify food hazards and extract supporting textual segments via BIO labeling.
- Implement a two-branch Hazard/Product architecture combining classification heads, segment extraction, and automatic loss weighting.
- Orchestrate a BIO annotation strategy using an ensemble of LLMs (Gemini, DeepSeek, Qwen) via OpenRouter, with soft labeling to handle justification subjectivity.
- Improve BERT baseline from 0.71 to 0.79 in F1 macro Hazard, with extractive explanations evaluated on sufficiency and completeness.

Dangerous Asteroid Detection (Python, scikit-learn) [🔗](#) Dec 2025

- Preprocess 90,000 NASA near-Earth objects through cleaning, feature selection, normalization, and class imbalance handling, then optimize models prioritizing recall to achieve approximately 99% on the « dangerous asteroid class ».
- Design an object-oriented Python architecture to compare six classifiers and generate visualizations improving model interpretability.

SKILLS & INTERESTS

- **Technical Skills:** Python, PyTorch, scikit-learn, Hugging Face, NLP / LLMs, RAG (LangChain, FAISS), REST API, Docker, Streamlit, Git.
- **Soft Skills:** Curiosity, critical thinking, rigor, adaptability.
- **Languages:** English (B2, TOEIC 795/990), Spanish (A2), Italian (A1).
- **Interests:** Regular running practice.