

Margot Berton | Data Scientist

LinkedIn · Portfolio website · GitHub

Montreal, Canada

Experience

Data Scientist

Nov 2024 - Present

BNP Paribas — New York, USA (Nov 2024 – Nov 2025) / Montreal, Canada (Nov 2025 - Present)

- Built an end-to-end RAG pipeline (LangChain, OpenSearch) for automated data extraction from business documents, achieving 80% accuracy (vs. 70% target) — deployed via Docker and Kubernetes.
- Prototyped a multimodal agentic system (RAG + visual analysis agents) for document understanding, targeting automation of 2,500 reports/year; evaluated using RAGAS, LLM-as-Judge, and Scikit-learn metrics, tracked with MLflow.
- Partnered with non-technical stakeholders to prioritize use cases and delivered actionable insights through data-driven dashboards (Streamlit, Matplotlib, Seaborn).

Computer Vision Intern

Feb 2024 – Aug 2024

MBDA — Le Plessis-Robinson, France

- Surveyed state-of-the-art explainability and uncertainty quantification methods (SHAP, LIME, GradCAM, Monte Carlo Dropout) and applied selected approaches to interpret an internal computer vision model, leveraging OpenCV and PyTorch for image processing pipelines.

Software Development Intern

Jun 2022 – Aug 2022

AiiNTENSE — Evry, France

- Developed a Python-based graphical interface for a medical pupillometer used in neurology and intensive care.
- Proactively reached out to a hospital team to understand identitovigilance requirements, leading to a dedicated workshop and the integration of patient safety measures into the tool.

Independent Research Project (TIPE)

2020–2021

Self-directed — Preparatory Class, La Martinière Monplaisir

- Independently reproduced a published mathematical model of cardiac cell renewal (cardiomyocyte turnover), based on a paper from *Interstices* whose author provided guidance — reaching a divergent model and identifying a potential alternative parameterization.
- Applied differential equations and numerical simulation (Python) to model biological dynamics, and critically compared results against the reference study.

Education

• Academic Exchange – AI Specialization

Aug 2023 – Dec 2023

Polytechnique Montréal — Montreal, Canada

• Master of Computer Science

2021 – 2024

Télécom SudParis — Evry, France (Institut Polytechnique de Paris)

Relevant coursework: Data Mining, NLP, Machine Learning, Algorithm Analysis and Design, Programming (Python, Java, C, SQL).

Activities

- Finalist (top 6/96 teams) in a startup pitch competition - proposed a Git-based video editing collaboration tool.
- Contributed to a student research project on fake news detection — built a dataset via Reddit web scraping and analyzed diffusion patterns of true vs. fake news.

Skills

- **Languages:** French (native), English (C1 — TOEIC 980/990)
- **ML stack:** pandas, PyTorch, scikit-learn, HuggingFace (transformers), MLFlow, GCP, OpenAI SDK, Langchain, MCP
- Developing expertise in **Responsible AI** — completing dedicated trainings and sharing insights through internal presentations at BNP Paribas.