

Baptiste Blouin

Aix-en-Provence, France
✉ BaptisteBlouin@proton.me
🌐 baptisteblouin.fr
in [baptiste-blouin](#)
🐙 [BaptisteBlouin](#)
French, English

Data Scientist & AI Engineer / NLP, LLM & Full-Stack

Professional Summary

Data Scientist and AI Engineer with a PhD in Computer Science, I combine published ML/NLP research expertise with strong hands-on engineering of AI products in production. From architecture to production, I design and build end-to-end AI products, autonomously. My SaaS platform illustrates this: a genuinely production-grade, multi-tenant LLM system where I cared as much about business value as operational robustness (LLM cost control and end-to-end security). My research background (7+ publications, international conferences, ERC grant) gives me a methodological rigor that I apply equally to designing technical architectures and solving complex problems.

Recent Projects

Aubier, Data & AI SaaS Platform

2025–Present

Full-Stack Design & Development

[Independent project, \[aubier-data.fr\]\(https://aubier-data.fr\)](#)

- **Multi-tenant SaaS platform** end-to-end: document extraction, RAG, text-to-SQL, analytics and predictions
- **Advanced RAG engine** (hybrid search pgvector + BM25, GraphRAG, RAPTOR) and **text-to-SQL** with SSE streaming
- **Full security**: JWT RS256, RBAC, tenant isolation, GDPR, Stripe billing
- **Stack**: FastAPI, PostgreSQL/pgvector, Celery, React 18, TypeScript, LangChain/LangGraph, Docker

baptisteblouin.fr: Portfolio & AI Agent

2026–Present

Production AI agent & public MCP server

baptisteblouin.fr

- **LangGraph agent in production** : multi-LLM fallback (Claude, Gemini, Mistral), **public MCP server**, evaluation 35/38

HistText: Large-Scale Text Analysis Platform

2023–2025

Full-Stack Research & Development

[Published: JMDMH 2024](#)

- **Text analysis platform** over 15 million historical documents (20 billion words), 200+ users
- **Stack**: Python, React, Rust, R, PostgreSQL, Apache Solr, Docker

Professional Experience

AI & Full-Stack Engineer, Aubier (*independent*)

Sep. 2025 – Present

End-to-end AI SaaS platform: architecture, async backend, multi-tenant security, React frontend.

Data Scientist & ML Engineer, Protisvalor, ENP-China Project, Aix-Marseille Université

Jan. 2023 – Aug. 2025

Lead developer of HistText (200+ users, ERC Proof of Concept grant); large-scale ML on historical texts; workshops and international conferences.

PhD Candidate & ML Scientist, Protisvalor, LIS / IrAsia, Aix-Marseille Université

Sep. 2019 – Dec. 2022

4 peer-reviewed publications during the PhD (7 in total); ERC Advanced grant; international conferences.

Skills

AI & LLM	RAG hybrid, GraphRAG, Text-to-SQL, AI Agents (LangGraph, MCP), Langfuse, litellm, Claude
NLP & ML	Multilingual NER, Transfer Learning, OCR, PyTorch, Hugging Face, spaCy, scikit-learn
Backend	Python, FastAPI (async), SQLAlchemy 2.0, Celery, Redis, PostgreSQL, pgvector
Frontend & Infra	React 18, TypeScript, D3.js, SaaS multi-tenant, RBAC, Docker, Cloudflare, GCP, n8n, CI/CD

Education

PhD in Computer Science (NLP / Machine Learning), Aix-Marseille Université

2022

Thesis: Event extraction from facsimiles of historical documents (Adv. Prof. Benoît Favre, ERC ENP-China project no. 788476)

MSc in Computer Science (AI/ML), Aix-Marseille Université

2019

BSc in Computer Science, Aix-Marseille Université

2016