Jean Savary

Senior Data Scientist & Backend Engineer Python

📧 jeansvy@gmail.com   🔗 linkedin.com/in/jean-savary-806681131

📍 Madrid, Spain (Remote + 1 visit/month in France)

Data Scientist & Backend Engineer with 6 years of freelance experience specialized in Python, Data Engineering, Machine Learning, and backend architecture. Expert in designing scalable APIs, data pipelines, and cloud deployment (AWS). Passionate about performance optimization and building robust data-driven solutions.

TECHNICAL SKILLS

Backend & Data
Python, Django, FastAPI, Flask, PostgreSQL, Redis, Celery, SQLAlchemy

Data Science & ML
Scikit-learn, Pandas, NumPy, TensorFlow, PyTorch, NLP, Time Series, MLOps

DevOps & Cloud
AWS (EC2, S3, Lambda, RDS), Docker, Kubernetes, CI/CD, Terraform, Linux

Databases & Tools
PostgreSQL, MongoDB, ArangoDB, Redis, BigQuery, ChromaDB, Airflow

PROFESSIONAL EXPERIENCE

Co-founder & Lead Data Engineer
April 2025 - Present

HoptyCV • SaaS Platform • Remote

- Development of a comprehensive AI-powered solution for automated skills portfolio generation
- FastAPI backend architecture with AI agents for CV parsing, skills extraction, and scoring
- Robust ETL pipeline for large-scale profile data processing and enrichment
- Dynamic template system with PDF generation and automatic adaptation to job offers
- LLM integration (OpenAI, Perplexity) for semantic analysis and skills matching
- Implementation of asynchronous task orchestration for batch CV processing

Stack: Python, FastAPI, PostgreSQL, Redis, Celery, Airflow, AWS (S3, Lambda, EC2), Docker, LLM APIs, Supabase

Co-founder & Backend Engineer
September 2024 - March 2025

Reevu • CRM for YouTube Creators • Remote

- Design and development of backend for YouTube video analytics platform
- ETL architecture for real-time ingestion of YouTube data (comments, metrics) via API
- ML models for sentiment analysis and insight extraction from comments
- Optimized database (PostgreSQL) for storage and fast queries on millions of comments
- NLP processing pipeline with vectorization and clustering for audience segmentation
- Webhooks and notification system for real-time alerts

Stack: Python, FastAPI, PostgreSQL, Redis, OpenAI API, Embeddings, NLP (NLTK, spaCy), AWS (EC2, RDS, S3), Docker, Celery

Chief Technology Officer (CTO)
April 2024 - December 2024

Seeqle.com • AI Candidate Sourcing • Paris (Remote)

- Complete overhaul of backend infrastructure and microservices architecture
- Design and implementation of optimized persona database for candidate search
- Airflow pipeline for automatic CV parsing, entity extraction, and skills scoring via LLM
- Cloud infrastructure with Railway and AWS for scalable deployment and monitoring
- Management of Backend (2-3 people) and Data (2-3 people) teams - architecture and code reviews
- Performance optimization: reduced CV processing time from 30s to 2s, API latency from 5s to 500ms
- Contribution to product strategy and technical roadmap definition

Stack: Python, FastAPI, PostgreSQL, Redis, Airflow, LLM APIs (OpenAI), Railway, AWS (EC2, RDS, S3), Docker, Kubernetes

Data Scientist & Backend Engineer
March 2024 - June 2024

Animaj • Media & Entertainment • Paris (Freelance)

- Automatic episode detection within compilation videos at large scale (~50K videos)
- Video processing pipeline: download with Pytube, frame extraction, vector encoding
- Custom embedding algorithm for visual content recognition and episode detection
- Distributed storage infrastructure with AWS S3 and ChromaDB vectorstore for fast similarity search
- Intelligent scoring system for automatic episode classification and grouping
- Deployment and orchestration on AWS EC2 for distributed processing

Stack: Python, OpenCV, Embeddings, ChromaDB, AWS (S3, EC2), Pytube, NumPy, Pandas, Scikit-learn

Team Leader | Data Science & Backend Python
December 2022 - September 2024

Diagoriente • Paris

- Management of a team of 4 data scientists/engineers on backend Python and ML projects
- Development of microservices architecture with FastAPI and Django
- Creation and maintenance of graph database (ArangoDB) for professional skills
- Design of REST APIs for ML-based recommendation engine

Data Scientist
September 2018 - March 2020

Evaneos • Paris Area

- Automatic NLP tagging tool for text data related to travel projects
- Ad-hoc analyses for sales, marketing, acquisition, and product teams
- Flask POC and development of internal reporting platform
- Online session scoring and pricing model (data mining, machine learning)
- Clustering analyses (customers, destinations) and causality
- Data team training: SQL, BigQuery, and Python

Stack: Python, Flask, BigQuery, Jupyter, Scikit-learn, NLP (TF-IDF, word2vec), Dask, Google Analytics

Other experiences:

Backend Engineer Python	MARBL	*July 2021 - October 2022*
Lead Teacher Data Science	Le Wagon	*June 2021 - December 2021*
Data Engineer	Stealth	*January 2021 - March 2022*
Data Scientist & Backend Engineer	Innov-Training	*July 2020 - December 2020*
CTO	Bouge	*June 2020 - March 2021*
Data Analyst Operations	MoovOne	*September 2021 - November 2021*

EDUCATION

Engineering Degree - Computer Science & Digital Technologies

IMT Atlantique • 2017 - 2020

Master of Science - Computer Science
École Polytechnique de Montréal • 2018

LANGUAGES

French (Native)   English (Professional)   Spanish (Professional)