

KONAN Amani Dieudonné

Senior Data Scientist | Principal AI Engineer | ML & Data Architect

Generative AI • MLOps • Real-Time ML Pipelines • Production AI Systems • Financial Services

Cocody Riviéra Faya, Abidjan, Côte d'Ivoire • (+225) 07 47 88 22 35 / 05 06 84 19 19

amani_dieudonne@yahoo.fr • [linkedin.com/in/amanidieudonnekonan](https://www.linkedin.com/in/amanidieudonnekonan)

Ivorian • Male • Born 12 March 1987, Didiévi, Côte d'Ivoire

EXECUTIVE SUMMARY

Senior Data Scientist and AI Engineer with 13+ years of international experience designing, building, and deploying production-level AI systems that deliver measurable business impact in financial services and multilateral institutions. Proven expertise in Generative AI (LLMs, RAG pipelines, multi-model orchestration with OpenAI, Claude, LLaMA, Gemini), AI Agents, Text-to-SQL, and vector database architectures. Track record building real-time ML pipelines using Spark, Kinesis, and Databricks, with hands-on MLOps across AWS (SageMaker, S3, Lambda, Redshift, Kinesis, EC2) and Databricks (Delta Lake, Unity Catalog, MLflow). Experienced in fraud detection, risk modelling, and predictive analytics for enterprise financial systems, including credit risk scoring and counterparty risk assessment at a pan-African trade finance institution. Deep proficiency in Python and SQL, with production deployments spanning ETL/ELT pipelines, CDC architectures, and CI/CD-automated ML workflows. Combines advanced econometric capability (ARIMA, VAR, VECM) with modern AI/ML engineering to drive data-informed strategy. Bilingual in English and French with extensive experience across African and Middle Eastern.

CORE COMPETENCIES

GENERATIVE AI & LLMs	MLOps & DATA ENGINEERING	PREDICTIVE ANALYTICS & RISK	DATA ARCHITECTURE & GOVERNANCE
<ul style="list-style-type: none">LLM Orchestration (OpenAI, Claude, LLaMA, Gemini)RAG Pipelines & Vector DatabasesAI Agents & Text-to-SQLLangChain & Prompt EngineeringNLP & Document IntelligenceMulti-LLM Platform DesignAWS Bedrock & SageMaker	<ul style="list-style-type: none">Real-Time ML Pipelines (Spark, Kinesis)MLOps (AWS SageMaker, Databricks MLflow)Databricks / Spark / Delta Lake / MLflowAWS (S3, SageMaker, Lambda, Redshift, Kinesis, EC2)ETL/ELT & CDC PipelinesCI/CD & Model Versioning (GitHub Actions)Feature Stores & Data Pipelines	<ul style="list-style-type: none">Fraud Detection & Anomaly DetectionRisk Modelling & Credit ScoringChurn Prediction & Customer AnalyticsTime Series Forecasting (ARIMA, VAR, VECM)Classification & Regression ModelsCausal Inference & A/B TestingImpact Evaluation & Econometrics	<ul style="list-style-type: none">Medallion & Lakehouse ArchitectureData Quality & Governance FrameworksMaster Data ManagementEnterprise Data Strategy & ERDsOracle, PostgreSQL, SQL ServerTableau / Power BI / StreamlitSAP Datasphere & DAMA-DMBOK

PROFESSIONAL EXPERIENCE

Manager, Data Scientist / Architect

Sep 2023 – Present

African Export-Import Bank (Afreximbank) - Cairo, Egypt

Lead teams driving production AI systems, enterprise data architecture, advanced analytics, and data science for afreximbank, delivering measurable business impact across risk, compliance, and operational intelligence.

Data Architecture & Strategy

- Designed and implemented end-to-end medallion architecture (Bronze/Silver/Gold) on Databricks with Delta Lake and Unity Catalog, establishing the enterprise data lakehouse foundation for all analytical and operational workloads.
- Architected Master Data Management (MDM) solutions consolidating party data across four enterprise systems (CRM, Core Banking, Treasury Management, ERP), creating unified golden records for institutional counterparties.
- Developed comprehensive data architecture blueprints including conceptual and logical data models, entity relationship diagrams, data dictionaries, and data flow documentation for mission-critical financial systems.
- Created the enterprise data quality framework with automated metrics scripted on Databricks, monitoring completeness, consistency, timeliness, and accuracy across all data domains.

-
- Led the selection and planned implementation of Ataccama as the enterprise data management platform, establishing data cataloging and metadata governance capabilities.

Data Engineering, Real-Time Pipelines & Cloud Infrastructure

- Built production-grade ETL/ELT pipelines extracting financial data from Oracle Finacle Core Banking System and Calypso Treasury Management System into PostgreSQL and Delta Lake targets.
- Implemented Change Data Capture (CDC) and real-time streaming capabilities for data synchronization between treasury management systems and the enterprise data warehouse, enabling near-real-time risk monitoring and analytics.
- Leveraged AWS services extensively: S3 for storage, Athena for serverless querying, Lambda for event-driven processing, RDS and Redshift for managed databases and warehousing, Kinesis for real-time streaming ML pipelines, SageMaker for model training and deployment, EC2 for compute – all orchestrated via boto3 SDK.
- Established GitHub-based CI/CD pipelines for all data engineering workflows, ensuring version-controlled, reproducible, and auditable deployments across development, staging, and production environments.
- Administered SAP Datasphere as the enterprise data warehousing platform, integrating structured and unstructured data sources across the organization.

Generative AI, MLOps & Advanced Analytics

- Designed and deployed a multi-LLM enterprise chat platform (OpenAI, Claude, LLaMA, Gemini via AWS Bedrock and Databricks) with AI Agents and Text-to-SQL capabilities; built RAG pipelines with vector database architectures (FAISS) for automated document intelligence and knowledge retrieval.
- Developed predictive analytics applications for real-time financial risk monitoring, counterparty risk scoring, and anomaly detection, delivering actionable intelligence to senior decision-makers.
- Automated and enhanced the African Commodity Index and internal commodity dashboards across 54 African countries, improving consistency, reproducibility, and analytical depth.
- Initiated the revamp of the methodological framework of the African Commodity Index to strengthen its analytical robustness and alignment with global best practices.
- Built end-to-end ML and econometric modelling pipelines (Python, PySpark) for trade flow prediction, regional integration patterns, and cross-country macroeconomic forecasting using IMF DOTS, UNCTAD, and WTO datasets, with MLOps workflows via Databricks MLflow and GitHub CI/CD.
- Improved research datasets through data quality processes, harmonisation, and methodological standardisation to support future econometric and policy analysis.

Data Engineer

Dec 2024 – May 2025

International Telecommunication Union (ITU) - Remote

- Designed and implemented an automated ETL pipeline on Databricks to process heterogeneous infrastructure datasets from six countries (.csv, .xlsx, .gpkg, .tiff), enabling scalable multi-format data ingestion.
- Developed a dynamic Data Quality (DQ) framework with rule-based checks (completeness, consistency, geographic integrity) and generated dual-format reports (.csv/.xlsx) with diagnostics and remediation guidance.
- Integrated backend logic into the CPP toolkit frontend, enabling users to upload data, trigger preprocessing, and download reports via a seamless web interface.
- Delivered YAML-based configuration for pipeline portability and no-code extensibility; coordinated full testing and deployment in shared compute environments.
- Produced a video walkthrough and technical handover package for long-term sustainability and knowledge transfer.

Statistician & Econometrician

Apr 2017 – Sep 2023

International Cocoa Organization (ICCO) - Abidjan, Côte d'Ivoire

- Designed and maintained a tailored statistical database application serving as the organization's central data architecture for cocoa market intelligence, streamlining reporting workflows and ensuring data consistency across member countries.
- Built forecasting models (ARIMA, VAR, VECM) to project global supply, demand, and pricing trends in the cocoa economy, directly informing strategic decisions for ICCO member countries and global stakeholders.
- Authored and contributed to the Monthly Market Reports and the Quarterly Bulletin of Cocoa Statistics, establishing robust data pipelines for timely and accurate dissemination of market insights.

- Led annual stock surveys across European warehouses to estimate regional and global end-of-season cocoa bean inventories, managing complex multi-site data collection and validation workflows.
- Collected, validated, and harmonized global cocoa market data from multiple international sources (futures markets, national statistics, trade data), implementing data quality standards and governance processes.
- Conducted econometric analyses of price volatility, supply–demand imbalances, and trade flow dynamics to produce evidence-based policy recommendations.

Researcher (Data Scientist / Statistician)

May 2023 – Aug 2023

United Nations Institute for Disarmament Research (UNIDIR) - Remote

- Designed multi-country survey data architectures (Nigeria, Colombia, Iraq, Lake Chad Basin) for the MEAC disarmament and reintegration research project, managing end-to-end data pipelines.
- Programmed and managed digital data collection via SurveyCTO with automated real-time quality control and validation rules; liaised with field teams on data consistency.
- Led data cleaning, harmonization, and multivariate statistical analysis pipelines using R and STATA, producing evidence-based insights for post-conflict transition outcomes.
- Delivered interactive dashboards and visualizations (Power BI) to support stakeholder engagement; ensured compliance with UN data governance and protection standards.

Data Analyst Consultant

Apr 2022 – Nov 2022

International Organization for Migration (IOM) - Remote

- Developed reproducible R scripts and markdown reports to analyze Migrant Health and Psychosocial Support (MHPSS) data across 8 West African countries, establishing standardized analytical pipelines.
- Conducted impact evaluation analyses on migration in Guinea, The Gambia, Nigeria, and Senegal, delivering evidence-based findings for programme improvement.
- Ensured analytical workflows were fully documented and automated for transparency and scalability across future assessments.

KoboToolBox Consultant

Dec 2016 – Feb 2017

United Nations High Commissioner for Refugees (UNHCR) - Abidjan, Côte d'Ivoire

- Designed and implemented a digital data collection system using KoboToolBox to support field operations and monitoring activities for refugee programmes.
- Trained field personnel on effective use of online data collection tools and ensured smooth transition from paper-based methods to digital workflows.
- Digitized all existing UNHCR Côte d'Ivoire data collection instruments and provided ongoing technical guidance to ensure data integrity and reporting consistency.

Monitoring & Evaluation Manager

Aug 2016 – Apr 2017

Barry-Callebaut (SACO) - Abidjan, Côte d'Ivoire

- Designed and managed the geotraceability data collection system for the Cocoa Horizons program, architecting data flows from field surveys through to analytical dashboards.
- Built and maintained master data systems for farmer organizations; defined and monitored KPIs through dashboards to inform strategic decision-making.
- Delivered a full M&E framework including data architecture, survey design, field execution, impact evaluation protocols, and progress tracking systems.
- Conducted fieldwork on cocoa demo plots and fermentation practices to evaluate agronomic innovations, combining field statistics with geospatial data collection.

Monitoring & Evaluation Officer

Nov 2015 – Jul 2016

Family Health International (FHI360) - Abidjan, Côte d'Ivoire

- Designed evaluation analysis plans and developed STATA-based scripts for automated data processing across multiple public health projects.

- Contributed to the development of four impact evaluation protocols and co-authored a peer-reviewed article on adolescent behavioral health.
- Drafted statistical sections of research protocols, study reports, and manuscripts; provided technical and statistical support for M&E activities.

Data Manager

Jun 2013 – Oct 2015

International Centre for Research in Agroforestry (ICRAF) - Abidjan, Côte d'Ivoire

- Designed and implemented M&E data architectures for agroforestry research projects; developed and maintained relational databases (Access, PostgreSQL) to monitor performance indicators.
- Conducted 10+ household studies (baselines, midterms, endlines) across West Africa; migrated data collection workflows from paper to digital systems using Open Data Kit.
- Automated data analysis and reporting pipelines using R, Excel, and Access; trained research teams in Kenya and Côte d'Ivoire on data quality best practices.
- Managed complex experimental and survey datasets; performed advanced statistical analyses using R and STATA for peer-reviewed research outputs.

Data Manager

Nov 2012 – Jun 2013

Programme ANRS Coopération Côte d'Ivoire (PAC-CI) - Abidjan, Côte d'Ivoire

- Harmonized 16 epidemiological datasets from 8 West African countries for the leDEA consortium, designing cross-country data integration architecture for HIV/AIDS research.
- Automated key project indicators reporting using custom-built SQL and VBA scripts, reducing manual effort and improving reproducibility.
- Conducted rigorous data quality assurance, validation, and consistency checks prior to submission to the regional hub in Bordeaux, France.

SELECTED ARCHITECTURAL & RESEARCH ACHIEVEMENTS

Enterprise Data Lakehouse Architecture

- Designed and deployed a full medallion architecture (Bronze → Silver → Gold) on Databricks, serving as the single source of truth for all institutional analytical workloads. Implemented Unity Catalog for centralized data governance, RBAC, and lineage tracking.

Mission-Critical Banking ETL Pipelines

- Built end-to-end production-grade ETL pipelines for Oracle Finacle CBS and Calypso TMS, with CDC capabilities ensuring near-real-time data synchronization for treasury and cash management analytics.

AI-Powered Data Quality & Anomaly Detection Framework

- Developed automated data quality and anomaly detection pipelines on Databricks measuring completeness, consistency, accuracy, and timeliness. Built ML-powered validation and fraud detection models integrating statistical anomaly detection, business rule enforcement, and real-time alerting.

Master Data Management Solution

- Architected an MDM solution consolidating party data from CRM, Core Banking (Finacle), Treasury (Calypso), and ERP systems into a unified golden record for the institution.

AI-Powered Enterprise Applications

- Designed NexusAI (multi-LLM enterprise platform with OpenAI, Claude, LLaMA, Gemini), RAG pipelines with vector databases (FAISS) for document intelligence, AI Agents with Text-to-SQL capabilities, and ALCO dashboard application (FastAPI + Tableau) providing real-time asset-liability and risk analytics.

EDUCATION & CERTIFICATIONS

2025	Professional Certificate in Data Engineering, AI & Data Science - Massachusetts Institute of Technology (MIT)
------	---

2021	MicroMasters in Data, Economics & Development Policy - Massachusetts Institute of Technology (MIT)
2020	Financial Modelling & Valuation Analyst (FMVA®) - Corporate Finance Institute (CFI)
2017	Master's in Statistics & Econometrics - Université Toulouse I Capitole, France
2015	Data Science Specialization Certificate - Johns Hopkins University
2012	Ingénieur Statisticien (Statistics & Information System Engineer) - ENSEA, Abidjan
2009	Bachelor in Statistics & Information Systems - ENSEA, Abidjan

TECHNICAL PROFICIENCY

Cloud Platforms	AWS (S3, SageMaker, Athena, Lambda, RDS, Redshift, Kinesis, EC2, Bedrock, boto3 SDK), SAP Datasphere, Databricks (Spark, Delta Lake, Unity Catalog, MLflow)
Data Engineering	ETL/ELT Pipeline Design & Development, CDC (Change Data Capture), Real-Time Streaming Pipelines, Medallion Architecture, Data Warehousing, Schema Evolution, Feature Engineering, Data Integration
Databases	Oracle (Finacle CBS), PostgreSQL, SQL Server, MySQL, SQLite, Microsoft Access
Programming	Python (PySpark, pandas, scikit-learn, LangChain, FastAPI, Streamlit, Dash, boto3), R, SQL, VBA, STATA, EViews
MLOps & DevOps	MLOps, GitHub Actions, Git Version Control, CI/CD Pipelines, Model Versioning, VS Code, JupyterHub, DataSpell
BI & Visualization	Tableau, Power BI, QGIS, ArcGIS, Streamlit Dashboards, R Shiny, Excel Dashboards
AI & ML	Generative AI (OpenAI, Claude, LLaMA, Gemini), LangChain, RAG Pipelines, Vector Databases (FAISS), AI Agents, Text-to-SQL, AWS Bedrock & SageMaker, NLP, Fraud Detection, Risk Modelling, Churn Prediction, Predictive Analytics, MLflow, Machine Learning,
Governance & Quality	Data Quality Frameworks, DAMA-DMBOK Principles, Data Cataloging (Ataccama), Metadata Management, Regulatory Compliance, RBAC
Econometrics	ARIMA, VAR, VECM, Panel Data Modelling, Causal Inference, Structural Break Analysis, Time Series Forecasting, Interrupted Time Series, GAMS
Survey & M&E	Survey Design & Sampling, SurveyCTO, ODK, Impact Evaluation, CPro, SPSS
Other	Postman, Mathematica, Scilab, Strapi, GitHub, Docker (basic)

LANGUAGES

French - Native Fluency | **English** - Full Professional Fluency

PROFESSIONAL ATTRIBUTES

Proven builder of production-level AI systems that deliver measurable business impact. Exceptional communicator capable of translating complex ML and AI concepts to executive stakeholders. Strategic thinker who aligns data science initiatives with business objectives and revenue outcomes. Self-driven, detail-oriented, and highly adaptable across diverse cultural and institutional contexts. Experienced in fully remote engagements for international clients, with strong cross-functional collaboration skills managing distributed teams across time zones.

REFERENCES

Dr. KOUAMÉ Christophe Regional Representative ICRAF c.kouame@cgiar.org +225 07 07 05 86 90	Dr. BONANNI Leonardo CEO and Founder SOURCEMAP leo@sourcemap.com +1 (617) 851 28 57	Mr. DIALLO Alpha Head of Administration & Finance UNHCR diallalp@unhcr.org +225 05 45 79 78 99
--	--	--