Coordonnées

33689294153 (Mobile)
xavier.vanausloos@ladonneei
ntelligente.com

www.linkedin.com/in/vanausloos
(LinkedIn)

Principales compétences

Modern Data Platform

OpenAI API

Multi-agent Systems

Languages

Français (Native or Bilingual)

Anglais (Full Professional)

Espagnol (Limited Working)

Russe (Elementary)

Certifications

Cloudera Data Plateform for CDH
Users

Google Cloud Platform Big Data and
Machine Learning

Apprentissage Statistique :
modélisation, prévision et data
mining

Talend for Big Data

Introduction to Recommender
Systems

Honors-Awards

Kaggle - Bosch Production Line
Performance - Prediction model
using XGBOOST for defect parts

Kaggle - Epilepsy seizure prediction

Xavier VAN AUSLOOS

Head of Data & AI | Data & AI Architect | Data Engineer
Aix-en-Provence, Provence-Alpes-Côte d'Azur, France

Résumé

Building scalable, future-proof data platforms with Airbyte, dbt, and
Dagster

With hands-on experience as a freelance Data Architect & Engineer,
I specialize in designing and deploying modern data platforms that
turn raw data into strategic assets. My recent work focuses on end-
to-end data orchestration—from ingestion (Airbyte) to transformation
(dbt) and workflow automation (Dagster)—to create scalable,
maintainable, and AI-ready pipelines.
I help organizations bridge the gap between data chaos and data-
driven decision-making by:
Architecting cloud-native and hybrid data lakes/lakehouses
(Iceberg, Delta, Databricks, Modern Data Platform).
Implementing DataOps best practices (CI/CD, monitoring,
governance) for reliable, self-service analytics.
Leading cross-functional teams (engineers, analysts, stakeholders)
to align data strategy with business goals.
Modernizing legacy systems with open-source tools (Airbyte, dbt,
Dagster) for cost-efficient, flexible solutions.

Tech Stack: Airbyte, dbt, Dagster, Spark, Databricks, Aws, Python,
DataOps.

Looking for a Head of Data, Data Architect, or Senior Data
Engineer?
Let's discuss how I can help you build, optimize, or scale your data
infrastructure—from POC to production.

DM me or comment below—always open to new challenges!

#DataArchitecture #ModernDataStack #DataEngineering #dbt
#Airbyte #Dagster #AI #ML #DataOps

———

Expérience

La Donnée Intelligente
CEO - Freelance Data Architect | Data Engineer
janvier 2016 - Present (10 ans 1 mois)
Marseille Area, France

Funder – Data Engineering & AI Consulting
Modern Data Stack | Scalable AI & Big Data Solutions

We specialize in end-to-end data engineering, AI/ML integration, and modern data platform consulting—helping businesses transform raw data into actionable insights and intelligent systems. Our expertise spans cloud-native architectures, real-time analytics, and generative AI, ensuring your data infrastructure is scalable, future-proof, and cost-efficient.

TechFirm Information Systems
Senior Data Architect
janvier 2025 - Present (1 an 1 mois)
Aix-en Provence

Data Architect for ADSN group (Notaires)
In charge of Data Transformation & Data Gouvernance project.

Design and setup of Modern Data Plateform (Dagster, Airbyte, Dbt, Datahub)
Consulting about Data Gouvernance & Data Catalog (DataHub, OpenMetadata)

Design and implement ML & IA solutions (OpenAI, LangChain, ChainLit, Chroma (vector DB),

Confidential Shipping start up
Data Engineer
septembre 2024 - décembre 2024 (4 mois)
Berlin, Germany

Provided support about Airflow setup and optimization and productionized a PySpark app

- Implemented a repo that runs a r-sync job with the Dag bucket as part of the CI/CD pipeline and DAGs optimization (Parallelizing Tasks, better code factorization, efficient scheduling with sensors)

- Refactored an existing PySpark notebook for production. Code Modularization: Notebooks tend to have inline, repetitive code. Parameterization: Hardcoded values make the notebook inflexible. Error Handling: Interactive code lacks error handling.
Optimize File Formats: Use efficient formats like Parquet, Version Control and CI/CD Integration, Unit Testing, Performance Tuning

Specific company names omitted due to confidentiality agreements

PLB Consultant
Data Engineering / MLOps instructor
avril 2023 - décembre 2024 (1 an 9 mois)
Paris, Île-de-France, France

Big Data & Data eng instructor : Cloudera, AWS, Databricks, Apache Spark, Scala, Java, Pyspark
Delivery of MLOPS training : MLFLOW, Apache SageMaker, CI/CD

Kolecto
Data Engineer / Data Ops Engineer
mai 2024 - octobre 2024 (6 mois)
Paris

Designed and built Modern Data Plateform (Dagster / Airbyte / Dbt) thanks to Terraform, Helm/K8S, AWS EKSStack: AWS, MDP, Terraform, Python, Github actions, EKS, Helm, K8S

Mirakl
Senior Data Engineer
octobre 2023 - mai 2024 (8 mois)
Paris, Île-de-France, France

Senior data engineer.
Development of Spark streaming / batch app in Scala, Java, Python
MLOPS with MLFlow for LLMs & ML models.
Stack : AWS / Databricks,
MLOPS with MLFLOW, GitHub actions

Institut Paoli-Calmettes
HealthCare Data warehouse product owner / Data consultant
août 2022 - septembre 2023 (1 an 2 mois)
Marseille, Provence-Alpes-Côte d'Azur, France

Consulting (master data management, data gouvernance, architecture of healthcare dwh, data modeling) and project management related to Healthcare Data Warehouse project.

https://www.institutpaolicalmettes.fr/decouvrir-ipc/la-strategie-data-a-lipc/

Confidential
AWS Senior Data Engineer
août 2022 - août 2023 (1 an 1 mois)
London, England, United Kingdom

Support the design of AWS data engineering solutions for a cloud migration project.

S3, Apache Spark, AWS Lambda, AWS Redshift, Amazon Managed Workflows for Apache Airflow, CI/CD Jenkins, GIT, SBT,Nexus, Maven, Java, Scala, Python, Docker, Databricks Delta Lake

SETRA Conseil
Apache Cassandra / Apache Spark instructor
juillet 2022 - août 2023 (1 an 2 mois)
Paris

Apache Cassandra & Apache Spark trainings

Zonova, Innovation and App. at the speed of your idea !
Consulting for Apache Cassandra & PowerBI
juillet 2022 - juillet 2022 (1 mois)
Lyon, Auvergne-Rhône-Alpes, France

Assist customer for tuning Apache Cassandra cluster for working with PowerBI

Quantmetry
Senior Data Engineer
avril 2022 - juillet 2022 (4 mois)
Paris, Île-de-France, France

Consulting & data pipelines development

Hortonworks - Cloudera CDP stack / Apache Atlas / Apache Hive / Apache Spark in Java (dedicated Java Spark framework) / AWS stack / S3 / PARQUET / DataLake

Zalando
Senior Data Engineer
septembre 2021 - avril 2022 (8 mois)
Berlin, Allemagne

For a large e-commerce company, developing pipelines using Apache Spark,Apache Flink, Apache Airflow in Scala / Java and Python Databricks, AWS Redshift, AWS S3, Databricks Delta Lake

Designing and implementing CI/CD process in Zalando Continous Delivery Plateform

Refactoring of large Apache Spark batch apps into Spark streaming apps using Databricks AutoLoader & Apache Spark Streaming & Apache Flink.

Nextra Partners
MLOps | Data Engineer
août 2021 - décembre 2021 (5 mois)
Marseille, Provence-Alpes-Côte d'Azur, France

Consulting in data science and data engineering for several Nextra Partners's customers

Confidential
Senior Data Engineer
septembre 2021 - octobre 2021 (2 mois)
Berlin, Germany

For a large e-commerce company, developing pipelines using Apache Spark,Apache Flink, Apache Airflow in Scala / Java and Python

SETRA Conseil
Consulting Apache Cassandra / Datastax, Apache Spark Datastax
juillet 2021 - septembre 2021 (3 mois)
Paris, Île-de-France, France

Consulting for several French IT services providers around Apache Cassandra and Apache Spark

VEOLYS BUILDING SERVICES
Data Scientist
décembre 2020 - septembre 2021 (10 mois)
St.-Raphaël, Provence-Alpes-Côte d'Azur, France

Creation of new data products using Data Science technics (clustering, PCA, hierarchical ascendant classification)

PLB Consultant
Big Data & Data Science Instructor
mai 2019 - septembre 2021 (2 ans 5 mois)
Levallois Perret

As freelance, in charge of Big Data trainings for professionals.

Cloudera Hortonworks, Spark, Hive, Scala, Flink, Deep Learning, Data Science, Apache Cassandra

Ecole nationale de la Statistique et de l'Analyse de l'Information
Associate Professor Big Data
novembre 2015 - septembre 2021 (5 ans 11 mois)
Région de Rennes, France

Lectures and labs about Big Data main tech: Hadoop, Cloudera, Aws, Spark, Flink

Pôle Emploi
Senior Data Engineer
février 2021 - août 2021 (7 mois)
Aix-en-Provence, Provence-Alpes-Côte d'Azur, France

Cloudera & Hortonworks Big Data Architect / Data Engineer (CDP, Hive, Hbase, Apache Phoenix, ElasticSearch, Spark, Scala, Kotlin, Kafka)

Grands acteurs webmarketing (confidentiel)
Data Engineer
janvier 2020 - janvier 2021 (1 an 1 mois)
Netherlands

Data Engineer (Cloudera/Hortonworks) Flink, Spark, Python, Scala, Airflow, Aws, Java, Cassandra

Smartpush
CTO - Data Architect | Data Engineer | Product owner
janvier 2019 - août 2020 (1 an 8 mois)
Marseille Area, France

AWS, Hortonworks, Cassandra, Python, Natural Langage Processing, Recommender Systems Data engineering with pyspark spark in java and scala

TRAXENS
Head of Data
janvier 2017 - janvier 2019 (2 ans 1 mois)
Région de Marseille, France

In charge of implementing data products solution and analytics architecture: choose of technologies (Cassandra/Spark/Scala/Python/R/Tableau) from idea to code and release with a team of 2/3 people.
Data anomalies detection, ETA prediction
Tech env: Datastax Entreprise/Cassandra, Python 3, Spark/Scala, Tableau, Pandas, etc.

Data Science algo used: LightGBM, XGBOOST, R Anomalies, ARIMA, SciKit, Hiera. Temp Memory

Université d'Aix-Marseille
Associate Professor Big Data / Data science Master Aix Marseilles School of Economics (AMSE)
septembre 2015 - septembre 2017 (2 ans 1 mois)

Students teaching in Big data and Data Science.

Python, Data Science, Big Data (Hadoop, Cloudera,Apache Spark, Hive etc.)

Sogeti
Big Data Architect, Data Engineer
janvier 2012 - janvier 2017 (5 ans 1 mois)
Aix-en-Provence, Provence-Alpes-Côte d'Azur, France

Consulting in Big Data & Data Science for Airbus Helicopters and CAC40 french companies
Tech: Data Lake, Hadoop, Data Mining, Data science, Python, Data engineering

La Donnée Intelligente
Consultant internet indépendant
janvier 2008 - novembre 2012 (4 ans 11 mois)
Région de Marseille, France

Web development (CakePHP, Zend et solutions open source Prestashop, J2EE Struts, Drupal, Spip, Symfony, Angular JS etc.)

Réalisations: www.kookbicycles.com, www.mareclamation.com, www.familleplanning.com, www.entreseniors.com, www.myrzo.com, www.voterepulic.com

CMA CGM
Consultant AMOA Invoicing & Product owner (SII for CMA CGM)
décembre 2006 - septembre 2012 (5 ans 10 mois)

Product owner : management of corrective and evolutive maintenance for worldwide CMA CGM invoicing module
AMOA: business requirements & framing, roadmap, acceptance strategies, project planning & follow up, change management, training, deployment strategies etc.

SII
IT Consultant

avril 2006 - septembre 2012 (6 ans 6 mois)

Le Tholonet (13)

IT consulting (web, java, Oracle, product owner)

HSBC

Web developer & project manager (JAVA / J2EE)

janvier 2002 - avril 2006 (4 ans 4 mois)

La Défense, Paris

Java/J2EE web developer in charge of on line banking (hsbc.fr,

hsbcprivatebank.com)

TotalEnergies

Network engineer and local IT manager

septembre 2000 - janvier 2002 (1 an 5 mois)

Assaluyeh, Iran

Responsible for local IT for all TotalFinaElf team in Assaluyeh, Iran (250 PC/10

servers/IT & Sat Telecom).

Management of a 15 iranian people team.

In charge of infrastructure (networks, servers, telcom), IT ops and users

support

Oki Printing Solutions

Product owner for printing solutions

avril 1999 - septembre 2000 (1 an 6 mois)

L'Hay Les Roses (94)

Functional and technical management for printing solutions.

Compaq International

L3 Support Engineer

septembre 1998 - mars 1999 (7 mois)

Dublin, Irlande

Level 3 network and app support for end users (in English and French)

ACTED

Project Manager

janvier 1997 - août 1998 (1 an 8 mois)

Kaboul, Afghanistan

———

Formation

Ecole Polytechnique Fédérale de Lausanne

Spécialisation Functional Programming in Scala / Spark, Information
Technology · (2017 - 2017)

Hortonworks Cloudera

Hortonworks Data Plateform (Big Data), Information
Technology · (2015 - 2015)

Massachusetts Institute of Technology

Data Science and Big Data Analytics, Information Technology · (2012 - 2012)

Université Grenoble Alpes

Master, Management Information Systems, General · (1992 - 1996)

IBM

Big Data with Hadoop stack by IBM, Big data · (2015 - 2015)