14 Years of experience

Demistifing the data realm

+33 6 74 29 08 54
mehdi@tazimehdi.com
tazimehdi.com

https://www.linkedin.com/in/mehditazi85

MEHDI TAZI
SENIOR (BIG) DATA ARCHITECT & TECH LEAD

EDUCATION AND CERTIFICATIONS

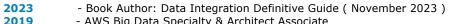|	
2023	- Book Author: Data Integration Definitive Guide (November 2023)
2019	- AWS Big Data Specialty & Architect Associate
2016	- PhD at Paris-EST in the Internet of Things (IoT) Virtualization field (In Progress)
	- Hadoop Hortonworks administrator certification
	- Fundamentals of Data Science
2015	- JAVA « OCAJ » Certification
2014	- Spring CORE Certification
	- Apache Cassandra for Java developers Certification
2013	- Scrum Master Official Certification - Scrum Alliance
	- MongoDB for DBAs and MongoDB for Java developers Certifications
2011	- Master's degree MIAGE (Business Informatics) in Engineering of Distributed Information Systems at (ESIAG Paris - France) (honors)
2010	- Engineering degree in Software Engineering Fez – Morocco (honors)

EXPERTISE

▪ *Technical & Architectural skills*
- o Defining a Data Strategy roadmap in alignment with business needs and top-level management strategy.
- o Consulting and Designing Data Platforms, LakeHouses, Datalakes, Datalabs and Modern data stack solutions: Organization, Architecture, Modeling, Governance, Security and Monitoring and Data Observability.
- o Participating in the development and maintenance of the IT architectural strategy and vision of enterprise architecture by continually evaluating trends in technology and business needs as well as selecting, designing and supporting implementation of new solutions for business transformation.
- o Designing and integrating existing systems, BI, analytical systems, data science systems with the (big)data platform.
- o Collecting business requirements, setting-up functional & technical solutions, then implementing monitoring & audits
- o Designing ingestion, exchange, logging/archiving, modeling and exposition big data layers/zones.
- o Setting-up BRP & BCP and Designing highly available, reactive, real-time, secure and governed solutions
- o Coding, Consulting and designing architectures distributed/Parallel scalable BigData/NoSQL solutions.
- o Data Engineering using Java/Scala/Python : AirFlow, Spark, DBT, AWS(EMR,GLUE,Athena,Lambda), Databricks…
- o Infrastructures & projects industrialization/Automation: DevOps, Infrastructure as code, MLOPS & Engineering stacks
- o Pre-sales and Writing technical specifications.
- o Technical Keynote speaker and trainer.

Storage	**DB:** Cassandra, MongoDB, DynamoDB, CosmosDB, HBase, Redis & RDBMS(Postgres, RDS …) **FILESYSTEMS/ObjectStorage:** HDFS, S3/Glacier/DeepArchive, ADLS2,… **FORMAT :** DELTA, PARQUET, CSV, XML, JSON, YAML, AVRO… **STREAM :** KAFKA, KINESIS…
Processing	Spark, HADOOP Technologies (Hive, SQOOP…), AWS (Glue, EMR, Databrew, DMS, Lambda, ECS, EKS …), DBT, Talend for big data, Dataiku, AZURE ADF…
Consumption & Exposition	Athena, Impala, Trino/Presto, GraphQL/REST API, PowerBI, QuickSight SageMaker, …
Automation	Docker, Airflow, Ansible, CloudFormation, Azure DevOPS, Jenkins, Terraform (in progress) …
Environments	**Infrastructure :** OnPremise, Cloud, Multi-Cloud, Hybride architectures **Operation Systems:** Windows, Unix, MacOS **Distributions/Providers:** AWS, ~~Azure~~, Cloudera, Hortonworks, Confluent/Aiven
Tools	Intellij IDEA, Visual Studio, Jupyter, Git, Maven/Sbt, Sonar, JIRA, QC, SNow, UML, Merise...
Prog languages	Python, Scala, Java/JEE, C#, HTML5/CSS3, SQL, CQL, JS

▪ *Project management skills*
- o Accompany teams setuping & onboarding their use cases
- o Coach, Lead, Estimate, Plan, Organize and monitor projects.
- o Animate meetings and motivate teams, Facilitate relations between and within teams.
- o Communicate on the goals & visions and write reports to the top management.
- o Prepare the releases, iterations and update the user stories and the backlog with the product owner.
- o Split the product into Use cases and prepare the releases, then Split Use cases into tasks and prepares the iterations.
- o Establishment of resources availability and of the production capacity tables.
- o Generating indicators like (BurnUp/Down charts, Team velocity, Tasks monitoring, Code Quality,Scope).
- o Identifications and risks management.

▪ *Tierce fields knowledge*
- o Stock Exchange, Real Estate, Travel, Web accessibility.
- o Business fields: Energy, Stock market, Retail banking, Business Travel, E-Commerce…

PROFESIONAL EXPERIENCES

- Lead (Big) Data Architect (application & infrastructure)

- Employer: IT INSIGHT – www.itinsight.fr

- Clients: Engie GEM, Engie IT **Paris – France** – www.engie.fr

- Job: I'm the lead big data architeict at Engie GEM, I've done many projects and studies as Big Data architect

3 year 10 months

- Missions :

- Responsible for Big Data architectures in a hybrid & multi-cloud context (OnPrem-Azure-AWS)
- Support & Consulting to promote the Big Data / Cloud approaches and patterns to develop high-performance use cases.
- Onboarding and support for teams in the implementation of their business use cases.
- Collaborate with the Business and IT to provide the overall guidance and structure for creating and promoting the Data approach and use cases
- Provide recommendations to the business in order to develop use cases using Data
- Collaboration with application architects, infrastructures, APIs, Security, DevOPS to guarantee the consistency of the data offer.
- Collaboration with the Engie Group for the evolution & improvement of data products (DataCatalog, Data services, organization ...).
- Implementation of a set of tools accelerating teams onboarding, code development, easing learning curve and promoting the big data / cloud culture: Code samples, Wiki / How to, Frameworks, Workshops, documents good practices and overall guidance.
- Collaboration for the continuous setup of the data strategy & vision through the evaluation and implementation of new technologies and approaches/patterns.
- Automation of standard data platform processes: FineOps & billing management, Monitoring, users & groups management...
- Study and implementation of several advanced subjects: API on Datalake, Governance, Security, Optimization, Automation & DevOPS ...
- Integration of the platform with third-party solutions: Collibra, Dataiku, PowereBI, Authentication, Splunk ...
- Industrialization of solutions thanks to DevOps and Infra as code
- Industrialization of log analysis for the automation of predictive maintenance & monitoring as well as the proposal of architectures meeting business requirements.
- Promoting & defining data guidance and structure for creating and promoting the Data projects and use cases.
- Participating in the development and maintenance of the IT architectural strategy and vision of enterprise architecture by continually evaluating trends in technology and business needs as well as selecting, designing and supporting implementation of new solutions for business transformation.
- Collaborating with Engie Group in evolving and enhancing the data products and data catalog for all Engie entities.
- Global data products offer (Data, SDK, Catalog, Billing ….)
- Cost and performance optimization of big data applications
- Management of development teams
- Monitoring and optimzing Billings
- Implementing an automatic solution for tracking un-optimizied big data / cloud solutions (cloud principes, queries respecting the partitioning …)

- **Tools:** Databricks, AWS, Spark, Glue services, S3, ATHENA, Kinesis / SNS, DMS, SageMaker, DynamoDB, Lambda, Docker, EC2, API GATEWAY, CloudFormation, CloudWatch, CloudTrails, AWS BATCH, EKS, Azure (DataFactory, ADLS, U-SQL, HDINSIGHT, EventHub), Scala /Python, Hadoop…

PROFESIONAL EXPERIENCES

1 year 10 months

- Big Data – NoSQL Architect (Application & Platform)
- Employer: IT INSIGHT – www.itinsight.fr
- Clients: AXA WORLD – TECH & **AXA – DATA INNOVATION LAB Paris – France –** www.axa.fr
- Job: I have done many projects and studies as Big Data architect and maintained the AXA Cluster
- Missions :

- Setting up diagrams of: technical architecture, high and low-level design, classes, sequence, component, de deployement and infrastructure & flows.
- Studying big data architectures, solutions and products (cloudera navigator, Kafka LDAP integration, Kafka Navigator integration ...)
- Setup & Conducts projects and PoCs
- Main lead on cloudera navigator which is used for data governance and access monitoring
- Main Designer & Architect of "SelfService" project which is a solution for automating operations on the big data platform (user creation, data ingestion,...), starting from the Silva ticketing Web Interface to the ansible scripts which deploys the new features & configuration on the platform. (Full DEVOPS + BigData capabilities)
- Administering the clusters and providing level three support: CDH5, HDFS, HIVE, SPARK, OOZIE, Flume…
- Coding and integrating tools for the big data platform with third party systems (RDMS, logging systems, ftp servers, LDAP, Kerberos ...) : CDH5, Kerberos, OpenLDAP, CentOS, postgresql ...
- Developing automation tools and scripts for deployment and operational procedures: Ansible, python ...
- Working on a secure and industrialized platform: Kerberos, LDAP, Certifications ....
- ❖ **Tools:** Hadoop Cloudera, Ansible, LDAP, Kerberos, HDFS, HIVE, KAFKA, SPARK, OOZIE, FLUME, PYTHON, JAVA, SILVA, HAPROXY, UNIX, PostgreSQL, Docker...

2 years

- Big Data – NoSQL Architect / Agile Coach (Application)
- Employer: OCTO Technology Paris – France – www.octo.com
- Clients: BNP PARIBAS, LA POSTE, COLISSIMO, EDF-DOAAT, ERDF, AL BARID BANK, SMCP, OCTO ACADEMY
- Job: I have done several projects mainly as a Big Data architect, and occasionally as an agile coach
- Missions:

- Identification of needs, use cases and Big Data perimeter; establishing roadmaps.
- MVPs and POCs Identification; Scooping, management and implementation of projects.
- BigData Consulting and accompaniment: Administration, Architecture, Governance and Security…
- Designing Big Data technical, conceptual, functional, organizational and methodological architectures.
- Designing Datalakes & Datalabs: Organization, Architecture, Modeling and Governance.
- Designing and integrating existing, BI and analytical systems with the big data.
- Designing ingestion, exchange, logging/archiving, modeling and exposition big data zones.
- Setting-up BRP & BCP and Designing highly available, reactive, real-time and secure solutions
- Design, Management and development of Big Data solutions.
- Scripts, infrastructures, deployments and configurations automation and industrialization.
- Pre-sales and Big Data training; Agile coaching
- ❖ **Tools**: Java/JEE, Scala, Hadoop, Spark, Kafka, Ansible, Docker, Cassandra, HBase, Hive, Sqoop, Parquet, Avro, HDFS, Ambari, Zookeeper, Conflient.io, Jupyter, Git, Jenkins, JHipster, Scrum…

6 years   **- Independent trainer: BigData Fundamentals, Hadoop, DevOPS, IoT, Scrum, Java...**

REFERENCES :

LANGUAGES
English: Fluent
French: Bilingual
Arabic: Mother tongue

PUBLICATIONS AND PUBLIC INTERVENTIONS
- Book Author : Data Integration Definitive Guide (November 2023)
- Co-Organizer of Casablanca **MongoDB User Group**
- Internet of Things Virtualization Poster for the "NIST Cloud Computing Forum & Workshop IX" at the National Institute of Standards and Technology – USA
- PhD paper about the Internet of things(IoT) virtualization focuses on « A Micro-Clouds sensors virtualization and composition approach for the Internet of Things »
- **« Games Programming »** Workshop animator & speaker for the JNJD at INPT Morocco
- Keynote speaker of technical and scrum management fields.
- Creating and participating to opensource projects: Chimera2D, MoovinParis,Itefacio.com
- Book Reviewer «**MongoDB for Java Developers**» edited by packtpub: goo.gl/YxskzV