

MEHDI TAZI

Architecte Data & IA | Directeur technique (CTO)

Stratégie & Conseil Data & IA | Plateformes sur site & Cloud

Architecture de solutions & d'entreprise | Ingénierie IA

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

Démystifier l'univers de la Data



Auteur

- Coauteur, The Definitive Guide to Data Integration — Packt, 2024.

[Voir le livre sur Amazon](#)

Formation et certifications

2019

- Certifications AWS Big Data – Specialty et Solutions Architect – Associate.

2016

- Recherche doctorale à Paris-Est sur la virtualisation de l'Internet des objets (IoT), en cours.
- Certification Hortonworks Hadoop Administrator.
- Fondamentaux de la Data Science.

2014-2015

- Certifications Java OCAJ et Spring Core.
- Certification Apache Cassandra for Java Developers.

2013

- Certified ScrumMaster — Scrum Alliance.
- Certifications MongoDB for DBAs et MongoDB for Java Developers.

2011

- Master MIAGE (informatique de gestion), Ingénierie des systèmes d'information distribués — ESIAG, Paris, France ; avec mention.

2010

- Diplôme d'ingénieur en Génie logiciel — Fès, Maroc ; avec mention.
- Certification Cisco CCNA — Rabat, Maroc ; avec mention.



Langues

Anglais — Courant.

Français — Bilingue.

Arabe — Langue maternelle.

Expertise technique et architecturale

Conception et modélisation de solutions

- Recueillir les besoins métiers ; définir les solutions fonctionnelles/techniques et les HLD/LLD des applications, infrastructures et flux d'intégration.
- Modéliser avec UML/Merise : vues classes, séquences, composants, déploiement, infrastructure et flux pour documenter les choix d'architecture.

Architecture de plateformes Data & IA

- Conseiller sur les plateformes Data, Lakehouses, Data Mesh, Data Lakes/Labs ; définir des blueprints alignant priorités métiers, ingestion, traitement et exposition.
- Concevoir les couches d'ingestion, traitement et exposition en médaillon Bronze/Silver/Gold ; définir zones de staging, échange, logs, archives et périmètres des workspaces.
- Planifier les migrations de plateformes/catalogues par coexistence et transition progressive : ex. AWS → Databricks, AWS Glue Data Catalog → Unity Catalog.

Architecture Cloud et ingénierie Data

- Concevoir des solutions sur site, hybrides et multicloud (AWS, Azure, Databricks) selon les besoins applicatifs et d'infrastructure.
- Conseiller, concevoir et développer des solutions Big Data/NoSQL évolutives, distribuées et parallèles.
- Développer les traitements Data en Java, Scala, Python avec Airflow, Spark, dbt, AWS (EMR, Glue, Athena, Lambda) et Databricks.

Conception d'applications IA

- Concevoir des plateformes IA et applications ML/LLM/agentiques ; définir les interactions entre modèles, outils et données et guider leur mise en œuvre.
- Intégrer systèmes, BI, analytique et Data Science aux plateformes Data ; définir accès gouvernés, calcul évolutif et supervision des workloads IA.

Ingénierie IA et industrialisation

- Avec le développement assisté par IA, traduire idées/besoins en spécifications logicielles, plans de réalisation et critères d'acceptation testables.
- Créer des harnesses (contexte réutilisable, instructions, validation automatisée) ; coder/tester avec l'IA, automatiser les pull requests avec revue humaine.
- Automatiser livraisons, exploitation, maintenance : DevOps, infrastructure as code, MLOps, stacks d'ingénierie ; améliorer le logiciel par supervision et retours.

Gouvernance, sécurité et fiabilité

- Établir catalogues techniques/métiers, lineage et dictionnaires partagés ; définir plans de qualité, contrôles et provisionnement des accès.
- Définir stratégies PII/données sensibles et modèles d'habilitation ; déployer outils d'observabilité, supervision et audits sécurité/architecture/techniques.
- Établir les PRA/PCA ; concevoir des solutions hautement disponibles, réactives, temps réel, sécurisées et gouvernées.

Architecture d'entreprise et leadership

Architecture d'entreprise et stratégie SI

- Évaluer les besoins métiers et tendances technologiques pour définir et maintenir la stratégie SI et la vision d'architecture d'entreprise ; sélectionner, concevoir et accompagner les solutions de transformation.
- Définir la stratégie Data & IA et les feuilles de route alignées sur les besoins métiers et les priorités de la direction.
- Aligner architectures applicative, d'infrastructure et Data sur les schémas directeurs cibles et feuilles de route de migration ; conseiller les équipes Data, Architecture et la DSI sur les choix et priorités de réalisation.

CTO, conseil et réponses aux appels d'offres

- Piloter les orientations technologiques, la stratégie Data & IA, le portefeuille de conseil et l'offre de services selon les besoins clients.
- Développer l'activité : identifier les besoins clients, qualifier les opportunités et cadrer les missions d'architecture et de conseil Data & IA.
- Piloter les réponses aux appels d'offres : propositions techniques, conception de solutions, approches de réalisation, estimations de charge et cahiers des charges techniques.
- Recruter et faire progresser les équipes techniques ; gérer les projets internes au forfait, du cadrage et chiffrage à l'affectation des équipes, la planification, le suivi et la recette.
- Guider les décisions d'architecture, superviser la qualité des solutions et accompagner les équipes sur les missions clients.
- Animer des conférences et formations techniques.

Modèle opérationnel cible et conseil

- Définir l'organisation de la plateforme, la structure des équipes, les rôles et responsabilités, les matrices de responsabilités et les backlogs de réalisation par équipe.
- Relier l'architecture cible à l'organisation chargée de son exploitation ; guider les équipes sur les priorités de réalisation et d'exploitation.

Gestion de projets

- Communiquer les objectifs et la vision ; préparer les rapports pour la direction.
- Évaluer la disponibilité des ressources et maintenir les plans de capacité.
- Accompagner et piloter les équipes ; estimer, planifier, organiser et suivre les projets.
- Décomposer les produits en cas d'usage pour planifier les releases, puis en tâches pour planifier les itérations.
- Préparer les releases et itérations ; maintenir les user stories et le backlog avec le product owner.
- Identifier et gérer les risques.
- Animer les réunions, mobiliser les équipes et renforcer la collaboration interne et entre équipes.
- Accompagner les équipes dans la mise en place et l'intégration de leurs cas d'usage.
- Produire les indicateurs de réalisation : burnup/burndown, vélocité, suivi des tâches, qualité du code et périmètre.

FinOps

- Définir et mettre en œuvre la gouvernance, les standards et les bonnes pratiques FinOps.
- Élaborer les prévisions de coûts et estimations financières.
- Participer aux comités d'architecture pour orienter les projets dès la conception vers des choix optimisant les coûts.
- Optimiser l'achat de services cloud.
- Optimiser l'infrastructure par le dimensionnement adapté des machines et clusters.
- Optimiser les coûts techniques par le bucketing, le partitionnement et le Z-ordering ; recommander des stacks légères et des arbres de décision pour le choix des technologies.
- Optimiser le stockage par niveaux de service et gestion du cycle de vie des données.
- Mettre en place des solutions FinOps pour identifier et supprimer les ressources orphelines, avec tagging automatique et base de gestion des configurations (CMDB).
- Automatiser les activités FinOps avec des outils internes : tableaux de bord Tableau et Power BI, et CMDB.
- Suivre les usages et dépenses ; produire un reporting régulier avec KPI.
- Animer les communautés FinOps et diffuser pratiques, culture et outils par webinaires, ateliers et événements.

Stack technologique

Plateformes Data & IA	Databricks · Plateformes Data natives AWS / Azure · Cloudera Private Cloud Base
Ingestion et messagerie	Azure Data Factory · AWS DMS · Amazon Kinesis · Azure Event Hubs · Kafka · Confluent / Aiven · Talend · Flume · Databricks Auto Loader · Lakeflow Connect · Amazon SNS · FTP
Traitement et transformation	Apache Spark · Hadoop · AWS Glue ETL · Amazon EMR · AWS Glue DataBrew · dbt · Lakeflow Spark Declarative Pipelines · Databricks Runtime / Photon
Orchestration et ordonnancement	Airflow · Lakeflow Jobs · AWS Batch
Stockage et bases de données	Amazon S3 · S3 Glacier / Deep Archive · ADLS Gen2 · HDFS · PostgreSQL · Amazon RDS · DynamoDB · Cosmos DB · Cassandra · MongoDB · HBase · Redis
Formats de tables et fichiers	Delta Lake · Parquet · Avro · CSV · XML · JSON · YAML
SQL et exposition des données	Databricks SQL / SQL Warehouses · Amazon Athena · Hive · Impala · Trino / Presto · Amazon API Gateway · REST / GraphQL APIs
Analytique et IA	Power BI · Tableau · Amazon QuickSight · Dataiku · Amazon SageMaker AI
Catalogues et partage	Unity Catalog · AWS Glue Data Catalog · Collibra · OpenSharing (Delta Sharing)
Sécurité et réseau	LDAP / OpenLDAP · Kerberos · HAProxy · AWS CloudTrail
Observabilité et exploitation	AWS CloudWatch · Splunk · Ambari · ZooKeeper · ServiceNow
Calcul et conteneurs	Amazon EC2 · AWS Lambda · Amazon EKS / ECS · Docker · Kubernetes
Livraison et automatisation	Git · Azure DevOps · Jenkins · Ansible · AWS CloudFormation · Terraform (en cours)
Langages	Python · SQL · Scala · Java / JEE · C# · CQL · JavaScript · HTML5 / CSS3
Outils de développement et qualité	Jupyter · IntelliJ IDEA · Visual Studio · Maven / sbt · Sonar · Jira · QC · JHipster · SILVA
Systèmes d'exploitation	Windows · Unix · macOS

Connaissances sectorielles

Secteurs : grande distribution et e-commerce, énergie, marchés financiers, banque de détail, immobilier, voyage et voyage d'affaires.

Expérience professionnelle

Decathlon

Architecte Data principal

5 ANS



Architecte Core Data et Data Lake.

- Accompagnement de la migration AWS vers Databricks et définition d'une stratégie de catalogue progressive : coexistence AWS Glue Catalog / **Unity Catalog**, puis adoption complète d'Unity Catalog.
- Conception de l'**architecture cible** Lakehouse / Data Mesh, des HLD/LLD et de la feuille de route stratégique : gouvernance, observabilité, données personnelles (PII) et modèle opérationnel cible.
- Contribution à l'évolution de la plateforme vers la Data & IA : exigences d'architecture pour les accès gouvernés aux données, le calcul évolutif et la supervision opérationnelle.
- Accompagnement de la mise en œuvre du socle de gestion et d'analyse des données : qualification des besoins, proposition de solutions et appui aux équipes pour réaliser les cas d'usage.
- Automatisation du provisionnement de la plateforme et de la gestion des accès ; mise en place de capacités de gestion et de protection des données personnelles (PII) et sensibles.
- Contribution à l'optimisation **FinOps** par les revues d'architecture et l'amélioration de l'efficacité technique.
- Contribution à l'extension de la plateforme dans plusieurs environnements cloud.
- Intégration de Databricks avec Collibra, les services AWS et des technologies tierces.

TECHNOLOGIES AWS · Databricks · Unity Catalog · dbt · Amazon EMR · AWS Glue · Amazon Athena · AWS Lambda · Amazon EKS / Kubernetes · Collibra

Expérience professionnelle

ENGIE

Architecte principal / Lead Big Data & Cloud — Applications & Infrastructure

4 ANS



ENGIE GEM et ENGIE IT | Paris, France | via IT INSIGHT

Études d'architecture et mise en œuvre de plateformes de données, d'applications et d'infrastructures cloud.

- Responsabilité de l'**architecture Big Data** sur site, Azure et AWS, dans des environnements hybrides et multicloud.
- Automatisation des contrôles des implémentations cloud inefficaces et des erreurs de partitionnement ; automatisation de la facturation, de la supervision et de l'administration des utilisateurs/groupes ; optimisation des coûts et des performances applicatives.
- Évolution de la **stratégie Data** et de la vision d'architecture d'entreprise ; analyse des technologies et des besoins métiers, conception de solutions et conseil sur les patterns d'architecture pour des cas d'usage performants.
- Pilotage des équipes de développement et intégration des cas d'usage ; collaboration avec les métiers et l'IT sur les opportunités Data, les orientations et les modalités de réalisation.
- Coordination des architectes applications, infrastructure, API, sécurité et DevOps pour assurer la cohérence de l'offre Data.
- Collaboration avec le groupe ENGIE sur les produits de données partagés, SDK, catalogues, services Data, facturation et organisation opérationnelle entre entités.
- Création d'exemples de code, wikis, guides pratiques, frameworks, ateliers et bonnes pratiques pour l'intégration, le développement et la montée en compétences Big Data / Cloud.
- Mise en place d'API sur le Data Lake, de la gouvernance, de la sécurité, de l'optimisation et de l'automatisation ; intégration de Collibra, Dataiku, Power BI, de l'authentification et de Splunk, avec DevOps et infrastructure as code.
- Industrialisation de l'analyse des logs pour la maintenance prédictive et la supervision ; proposition d'architectures alignées sur les exigences métiers.

TECHNOLOGIES Databricks · AWS · Apache Spark · AWS Glue · Amazon S3 · Amazon Athena · Amazon Kinesis · Amazon SNS · AWS DMS · Amazon SageMaker · Amazon DynamoDB · AWS Lambda · Docker · Amazon EC2 · Amazon API Gateway · AWS CloudFormation · Amazon CloudWatch · AWS CloudTrail · AWS Batch · Amazon EKS · Azure · Azure Data Factory · Azure Data Lake Storage · U-SQL · Azure HDInsight · Azure Event Hubs · Scala · Python · Hadoop · Collibra · Dataiku · Power BI · Splunk

AXA

Architecte Big Data & NoSQL — Applications & Plateforme

2 ANS



AXA World Tech et Data Innovation Lab | via IT INSIGHT

Projets et études d'architecture Big Data, et maintien du cluster Big Data d'AXA.

- Production des HLD/LLD et de la documentation d'architecture technique : diagrammes de classes, séquences, composants, déploiement, infrastructure et flux ; lancement et pilotage de projets et de preuves de concept.
- Évaluation de solutions Big Data et de l'intégration Kafka-Navigator ; pilotage de Cloudera Navigator pour la gouvernance et le suivi des accès.
- Conception de l'architecture SelfService pour automatiser les opérations de la plateforme, du ticketing aux scripts Ansible ; automatisation des déploiements et des opérations avec Ansible et Python.
- Administration des clusters et support de niveau 3 sur CDH5, HDFS, Hive, Spark et Oozie.
- Intégration des outils de la plateforme avec les bases relationnelles, les logs, FTP, LDAP et Kerberos, avec CDH5, OpenLDAP, CentOS et PostgreSQL.

TECHNOLOGIES Cloudera Hadoop · Ansible · LDAP · Kerberos · HDFS · Hive · Kafka · Spark · Oozie · Flume · Python · Java · SILVA · HAProxy · Unix · PostgreSQL · Docker · Cloudera Navigator · CDH5 · OpenLDAP · CentOS

Expérience professionnelle

BI-NEWVISION

Directeur technique (CTO) & Architecte Data principal

7 ANS



Direction technique et conseil Data & IA

Piloter la stratégie technologique et l'offre de conseil Data & IA de BI-NEWVISION, en associant direction de l'entreprise et missions de conseil en architecture.

- Définir et faire évoluer la **stratégie Data & IA**, le portefeuille de conseil et l'offre de services en alignant les capacités techniques sur les besoins clients.
- Développer de nouvelles activités en qualifiant les opportunités et en cadrant les missions d'architecture et de conseil Data & IA.
- Piloter les réponses aux appels d'offres : conception de solutions, propositions techniques, approches de réalisation et estimations de charge.
- Définir les orientations techniques des missions, guider les décisions d'architecture et superviser la qualité des solutions.
- Recruter, piloter et accompagner les équipes ; gérer les projets internes au forfait, coordonner les priorités et assurer le conseil en architecture sur les missions clients.

OCTO Technology

Architecte Big Data & NoSQL / Coach Agile — Applications

2 ANS



BNP PARIBAS, LA POSTE, COLISSIMO, EDF-DOAAT, ERDF, AL BARID BANK, SMCP, OCTO ACADEMY

Missions de conseil en tant qu'architecte Big Data et, ponctuellement, coach Agile.

- Cadrage des besoins métiers et des cas d'usage Big Data, définition des feuilles de route et des MVP/preuves de concept, et pilotage des projets de mise en œuvre.
- Conseil opérationnel en administration, architecture, gouvernance et sécurité ; conception d'architectures techniques, conceptuelles, fonctionnelles, organisationnelles et méthodologiques.
- Conception de Data Lakes et Data Labs : organisation, modélisation et gouvernance ; intégration des systèmes existants, de la BI et de l'analytique aux plateformes Big Data.
- Conception des couches d'ingestion, d'échange, de journalisation, d'archivage, de modélisation et d'exposition des données.
- Mise en place de plans de reprise et de continuité d'activité (PRA/PCA) ; conception de solutions hautement disponibles, réactives, temps réel et sécurisées.
- Conception, pilotage et développement de solutions Big Data ; automatisation des scripts, de l'infrastructure, des déploiements et de la configuration.
- Appui à l'avant-vente, animation de formations Big Data et coaching Agile/Scrum.

TECHNOLOGIES Java / JEE · Scala · Hadoop · Spark · Kafka · Ansible · Docker · Cassandra · HBase · Hive · Sqoop · Parquet · Avro · HDFS · Ambari · ZooKeeper · Confluent.io · Jupyter · Git · Jenkins · JHipster

Références et publications

Références

Thierry GRIMA

Directeur Analytics Groupe

ENGIE & AXA

linkedin.com/in/thierrygrima/

Raphael MANSUY

Architecte Data

DECATHLON

linkedin.com/in/raphaelmansuy/

Publications et interventions

- Coauteur, The Definitive Guide to Data Integration — Packt, 2024.

[Voir le livre sur Amazon](#)

- Relecteur technique de MongoDB for Java Developers — Packt Publishing.

goo.gl/YxskzV

- Coorganisateur du Casablanca MongoDB User Group.
- Poster Internet of Things Virtualization présenté au NIST Cloud Computing Forum & Workshop IX — National Institute of Standards and Technology, États-Unis.
- Article de recherche doctorale sur la virtualisation de l'IoT : « A Micro-Clouds Sensors Virtualization and Composition Approach for the Internet of Things. »
- Animateur et intervenant de l'atelier « Games Programming » aux JNJD, INPT, Maroc.
- Conférencier sur des sujets techniques et le management Scrum.
- Créateur et contributeur de projets open source : Chimera2D, MoovinParis et ltefacio.com.
- Auteur d'articles techniques sur Medium.

medium.com/@mehdi.tazi