

Omar SAYAH

126 Avenue Roger Salengro
92290 Chatenay-Malabry
E-mail : sayahomar@outlook.com
Tel : 06.68.18.24.67

Data Architect Senior Data Engineer & Artificial intelligence

FORMATIONS

2013 Paris	Master 2 en Informatique option : Intelligence Artificielle et décision Université Pierre et Marie Curie, Paris Sorbonne.
2011 Versailles	Master 2 en Informatique option : Cryptographie et sécurité des Systèmes Informatique. Université de Versailles, Paris Saclay
2010 Bejaia	Ingénieur en Informatique option : systèmes parallèles et distribués Université UAMB de Bejaia
2006 Bejaia	Année préparatoire Sciences Exactes, Technologies et Informatique Université UAMB de Bejaia

COMPÉTENCES

Big data	Pyspark, Spark SQL, Deltalake, Databricks Data Engineer, Synapse,
Intelligence Artificielle	Weka, KXEN, Mahout, Azure AI Foundry, Databricks Machine Learning
Business Intelligence	MSBI (SSIS, SSAS, SSRS), MDX, DAX, Power BI, Master Data Services, DBT, OLTP, OLAP, Snowflake,
Bases de données	SQL Server, PostgreSQL, Oracle, SQL datawarehouse
Cloud	Azure Datafactory, Azure Logic Apps, Azure Function, Azure Storage, Azure Purview, Azure DevOps, Azure KV, Azure API Management, Azure monitor
Langages de programmation	Python, Scala, Java, C#.net, SQL, Yaml, Shell, power Shell
Ordonnancement	Control M, One automation, Airflow
Méthodologies	Méthode Agile
Architecture et Modélisation	MERISE, UML, Kimball, Data Vault, Entités Relations, architecture en médaillon, data hub
EDI/outils	VS Code, IntelliJ IDEA, Sublime Texte, Maven, SonarQube, Dax Studio, SSMS, Git, Datadog,

LANGUES

Anglais : Courant

EXPÉRIENCES PROFESSIONNELLES

Juin 2024 – à aujourd’hui : UNEO (Mutuelle santé)

Data Architect / Senior Data engineer

Projets :

- 1- Architecture, développement et optimisation des data flow et des rapports Power BI**
- 2- Mise en place d’un système de calcul des cotisations des salariés à partir des fichiers DSN**

Tâches :

- Etude et optimisation des data flow et des rapports Power BI existants,
- Proposition de nouvelles architectures et des data model pour les nouveaux projets,
- Développement des modules en T-SQL, Pyspark et Python,
- Développement des pipelines d’alimentation de données Datafactory sous Azure Synapse,
- Optimisation et développement de rapports Power BI,
- CICD et automatisation des releases avec Azure Devops,
- Rédaction des documentations techniques

Technologies : Azure Synapse, Pyspark, Spark SQL, Azure Datalake, Python 3.12, Power BI, Azure Devops, Yaml

Novembre 2022 – à Mai 2024 : Neuflyze - Groupe ABN AMRO (Banque privée)

Senior Data Engineer

Projets :

- 1- Développement d’un DATAHUB pour centraliser les données de la banque**

Tâches :

- Modélisation des données dans le Datalake à partir des données source, et création des tables Delta dans les différentes zones (architecture en médaillon),
- Développement des flux de transformation des données ETL en Pyspark et en SQL avec DBT,
- Développement des pipelines d’alimentation des données avec DBT, Datafactory et Air flow,
- Optimisation des flux et des pipelines de données existants,
- Développement des pipelines CI/CD,
- Review des codes,
- Modélisation et développement de rapports et de tableaux de bord Power BI
- Accompagnement de l’équipe recette pendant les phases de recette
- Monitoring de la production et résolution des incidents,
- Rédaction des documentations techniques

Technologies: Azure Databricks, Azure Datafactory, DBT, Airflow, Python, Pyspark, Spark SQL, Azure Datalake, Power BI, Azure Devops, Yaml, Data vault

Octobre 2021 - Octobre 2022 : TF1 (Audiovisuel)

Techlead data engineer et Intelligence Artificielle

Projets :

- 1- Suivi, maintenance et évolution des data flow dans le DATAHUB
- 2- Suivi et développement d'une solution d'optimisation des campagnes publicitaires TVS
- 3- Suivi et développement d'une solution pour l'optimisation des GRP en utilisant l'IA

Tâches :

- Alimentation du backlog par des US techniques,
- Accompagnement et encadrement d'une équipe de 6 développeurs,
- Participation à des ateliers techniques et d'échanges avec les architectes,
- Modélisation, et proposition de nouvelles architectures et de nouveaux data model,
- Optimisation des flux existants,
- Review et validation des codes,
- Participation aux développements des data flow,
- Accompagnement des nouveaux arrivants,
- Planification et élaboration de la roadmap technique,
- Veille technologique, et échange avec les experts Databricks et Microsoft sur nos problématiques

Technologies : Azure Databricks, Azure SQL Server, Azure Data Factory, Spark Scala, Pyspark, TSQL, Power BI, Azure Function, Azure Devops, Yaml, Azure Datalake gen2, API REST

Aout 2019 – Octobre 2021 : VINCI Énergies (Télécom et énergie)

Data Architect/ Senior Data Engineer

Projets :

- 1- Réalisation des POC, modélisation et élaboration des architectures Data et Business Intelligence,
- 2- Suivi et développement d'un système d'échange de données informatisées (EDI),
- 3- Suivi et développement d'un ensemble de projets pour la gestion des interventions,

Tâches :

- Études de faisabilité technique et réalisation des POC,
- Proposition et chiffrage de nouvelles architectures et solutions techniques,
- Rédaction des STD spécifications techniques détaillées et alimentation,
- Suivi et participation au développement des nouveaux projets,
- Suivi et accompagnement des développeurs juniors,
- Review et validation des codes,
- Modélisation et développement des rapports Power BI
- Développement et configuration des pipelines CI\CD Azure Devops, Rédaction des documentations techniques, et validation des livrables, Formation des utilisateurs

Technologies : Azure Databricks, Pyspark, Spark SQL, Azure SQL Server, Azure Data Factory, Power BI, DAX, Azure Logic Apps, API REST, Azure Devops

Janvier 2017 à Mai 2019 : ENGIE TRADING (Banque et énergie)

Senior data engineer

Projets :

- 1- Pilotage et MCO des datawarehouses et des systèmes décisionnels,
- 2- Pilotage et MCO des applications IT Finance et back office,
- 3- Migration et monter de version des applications,
- 4- Support : En charge du suivi de la qualité et de la disponibilité des données quotidiennement

Tâches :

- Recueil des besoins métiers pour les projets d'évolution avec l'aide des BA,
- Maintenance et évolution des datawarehouses et des bases de données,
- Maintenance des applications C#.NET,
- Optimisation des procédures stockées et des scripts T-SQL et PL-SQL,
- Migration de bases de données SQL serveur vers Azure SQL Server,
- Migration des packages SSIS, cubes SSAS et des rapports SSRS de MSBI 2008 vers 2012/2016,
- POC Big data avec Azure HDInseight pour la reconsolidation des données comptables,
- Optimisation et automatisation des processus de la clôture comptable,
- Support technico-fonctionnel et suivi de la production,
- Gestion des relevés bancaires (CODA, MT940, AFB120) et des paiements pour la trésorerie,
- Intégration des données (CRE, Couverture, Échéanciers...) pour les équipes comptables,

Technologies : T-SQL, PL/SQL, MSBI 2012/2016, MDX, C#.NET, One Automation, Spark, python, Hive, HDInseight ADF gen1, VSTS, blob storage

Octobre 2013 à Décembre 2016 – Crédit Agricole CIB (Banque)

Data engineer

Projets :

- 1- Développement d'une chaîne décisionnelle from-scratch pour le calcul d'indicateurs de risque (VAR, SVAR, PNL...) CVA et HEDGE,
- 2- Migration et optimisation de la chaîne décisionnelle de calcul d'indicateurs de risque CREDIT,
- 3- Support et MCO : Assurer la disponibilité, la cohérence et la qualité des données dans les deux applications SAFIR CVA et SAFIR CREDIT,

Tâches :

- Modélisation et développement des nouveaux Datawarehouses avec l'aide du techlead, Modélisation et développement des packages SSIS et des cubes SSAS,
- Conception d'une chaîne Control M pour l'intégration automatique des données, Documentation des différentes phases du projet,
- Migration des packages SSIS 2005 et des cubes SSAS existants vers MSBI 2012, Optimisation des procédures stockées et des vues,
- Parallèle Run et tests de non-régression,
- Conception de nouvelles chaînes Control M pour les projets existants,
- Suivi quotidien des intégrations des données dans les Datawarehouses et les cubes, Analyse des écarts remontés par le business,

- Réalisation des stress tests, des stress tests EBA et de recalibrage avec le business, Accompagnement des utilisateurs,

Technologies : MSBI 2012 (SSAS, SSIS), T-SQL, C#.NET, TFS, Control M, Power Shell, Python, Datawarehouse.

Avril 2013 à Septembre 2013 : Aéroport de Paris (Transport)

Stagiaire développeur BI, Base de données,

Projet :

1. Migration, optimisation et monter de version d'une chaine BI existante

Tâches :

- Migration du Datawarehouse de : Oracle 10g vers Oracle 11g et optimisation /PLSQL,
- Migration des jobs ETL BODI vers BODS,
- Migration des scripts Shell Unix vers Power shell,
- Test de non-régression et étude et analyse de la qualité des données après migration,
- Mise à jour de la documentation,
- Participation à la mise en PROD

Technologies : ORACLE 10G, ORACLE 11G, BUSINESS OBJECT DATA SERVICES, POWER SHELL, TOAD,

01/04/2011 au 31/09/2012 : Laboratoire PRISM/BluWan (Télécom)

Ingénieur R&D

Projet :

1- Étude de la sécurité des données dans le réseau sans fils Bluwan

Tâches :

- Études et recueil des besoins client,
- Études et analyse d'un ensemble d'articles scientifiques,
- État de l'art sur les différentes attaques qui peuvent dysfonctionner un réseau sans fils,
- Proposition de plusieurs solutions de sécurité, et rédaction des spécifications techniques
- Simuler un ensemble de solutions retenues avec le client,
- Analyse et présentation des résultats obtenus,
- Rédaction de la documentation technique des solutions retenues.

Technologies : C++, NS3, LINUX, SHELL, 802.11N, PROTOCOLES DE ROUTAGES,

Stagiaire R&D

Projet :

1- Proposition d'une nouvelle solution de load balancing dans les serveurs web

Tâches :

- État de l'arts sur les différentes solutions load balancing dans un data center (basées : DNS, Switch, software...),
- État de l'art sur les différentes politiques de remplacement des données dans les mémoires cache,
- Conception et modélisation d'une nouvelle solution d'équilibrage de charge basée DNS avec mémoire cache appelée DLB-TSC,
- Implémentation de la solution dans un simulateur en langage Java, et simulation en utilisant un fichier log d'un serveur web réel,
- Rédaction et interprétation des résultats obtenus,
- Publication d'un article dans un journal scientifique (voir le lien ci-dessous). <http://ieeexplore.ieee.org/document/6970100/> .

Technologies : JAVA, LINUX, LOAD BALANCING, MEMOIRE CACHE, SIMULATION, ALGORITHMIQUE, QUALITY OF SERVICES.