|  |
| --- |
| Muhammad Ehsan  Certified Data Architect | Manager Data Engineering & Advance ANALYTICS  Phone: +1 571 554 4040  Email: ehsan4u@gmail.com  LinkedIn: linkedin.com/in/ehsanullah |

|  |
| --- |
| Results-driven data engineering **leader** with extensive experience in designing and implementing scalable data architectures, data governance **frameworks**, and advanced analytics solutions. Proven track record in driving digital **transformation** across industries such as aviation, manufacturing, healthcare, and **government**. Expertise in cloud platforms (GCP, Azure, AWS), big data technologies, and automation tools. Passionate about leveraging data to drive business value and innovation. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | | | | |
| Leadership Cross-functional team management, agile project delivery, and stakeholder collaboration Skills **Big Data Tools:** Apache Spark, Hadoop (CDP/CDH/HDP) , Hive, Kafka, Airflow, Apache NIFI , DBT, Databricks (Delta Lake), Spark, Hortonworks (Ambari), Pig, Flume, Sqoop, TDCH, Knox, Kerberos, Ranger  **Data Modeling:** Erwin Data Modeler  **Cloud Data Platforms:** GCP (BigQuery, Dataflow, Dataproc, Google Cloud Storage), Azure (Azure Data Lake Storage (ADLS), Databricks, Azure Synapse, Data Factory), AWS (Databricks, S3, Glue, Redshift), Snowflake  PostgreSQL, MySQL, SQL Server, Redshift  **Data Engineering**: ETL/ELT pipelines, data lake/Lakehouse architectures, real-time data processing  **Data Governance** & **Cataloging**: Data quality frameworks, metadata management, data lineage, and compliance (GDPR, SOX), Unity Catalog (Metadata & Governance), Microsoft Purview, Dataplex  **Machine Learning & AI:** Predictive modeling, MLOps, TensorFlow, PyTorch, Scikit-learn **Programming:** Python, PySpark, SQL, Bash, Java, C#  Data Visualization: Power BI, Tableau, SAS  Automation & DevOps: Ansible, Puppet, Docker, Kubernetes, CICD  **Schedulers:** Informatica, Control-M, Autosys, Cloud Composer CERTIFICATIONSAlteryx Certified – Designer CORE **Alteryx** Server Administrator  **Tableau** Data Analyst **Tableau** Server Associate Data Engineer Associate (Databricks)Oracle Big Data Appliance – AdministrationCloudera Certified Data Analyst (CCA 159)Cloudera Certified Hadoop & Spark Administrator (CCA 131)PMP Certified (PMI ID: 6307991)Certified Ethical Hacker (2013)Six Sigma **Green Belt** & **Yellow Belt** CCISP (2014)  ITIL Foundation v3 |  | Experience | | | | |
|  | SENIOR Manager, Finance Business Intelligence (QATAR AIRWAYS) | | | 2023-Present | |
|  | Lead the development and implementation of enterprise data analytics solutions to optimize operational efficiencies and support data-driven decision-making.  Spearhead data governance and quality frameworks, ensuring compliance with industry standards and enhancing data accessibility.  Manage migration of on-premise data warehouses to a scalable, high-performance data warehouse cloud-based solutions (Azure, Google BigQuery), improving scalability, ensuring optimal storage, query performance, cost efficiency and performance.  Develop and deploy 14+ analytical data products for finance, supply chain, and flight operations, enabling real-time insights for stakeholders.  Automate ETL pipelines using SQL and Apache Airflow, processing over 5 million records daily with 99.9% accuracy.  Enhance database performance by 45% through query optimization and index tuning.  Designed robust, reusable and scalable data solutions and pipeline frameworks to automate ingestion, processing and delivery of batch and real-time streaming data.  Built and maintained conceptual, logical and physical data models for transactional, BI and analytical platforms. Played an active role in all stages of a data solution life cycle: acquisition, ingestion, curation, publication, quality and metadata to ensure consistency with data management standards and best practices. Oversee a team of 20+ engineers that build and manage large-scale developed data models and schema design optimized for analytical workloads, including partitioning and clustering strategies to enhance query efficiency  - Perform in-depth code reviews and manage GitHub repositories that contain data pipelining and data transformation code  - Architect data warehousing solutions to meet the needs of analysts, including building views to expose real-time data in BigQuery, Integrated Google Cloud Storage (GCS), Pub/Sub, Cloud Functions, and Dataflow to enable seamless ingestion of structured and semi-structured data on a regular basis, securing sensitive data using dynamic data masking policies, and maintaining a detailed data catalog  - Promote the use of data throughout the company through BigQuery, Tableau, Alteryx training sessions to encourage more data-driven decision making  - Manage team financials and budgets, build relationships with external software vendors, and identify new technologies that can further benefit the organization  - Present team achievements to executive stakeholders and share key learnings at external conferences | | | | |
|  | Manager, Data Engineering & Advanced Analytics (Pratt & Whitney) | | | 2020-2023 | |
|  | Directed data engineering **initiatives** across multiple business units, focusing on data governance, quality, and security. Led the migration of a Cloudera-based data lake to Databricks, data assessment, strategy planning, improving data processing efficiency, execution and scalability.  Designed and implemented data platforms for predictive maintenance, reducing downtime and saving millions in operational costs.  Developed advanced analytics solutions for supply chain optimization, reducing production delays and improving forecasting accuracy. Established data governance frameworks to ensure compliance with GDPR and other regulatory requirements. Managed the migration of a Cloudera data lake to Databricks (Delta Lake).  Implemented data governance frameworks to ensure data quality, lineage, and compliance with industry regulations.  Supported AI/ML initiatives by building ETL pipelines and deploying machine learning models on AWS and Azure. ETL data from SAP ERP modules (Finance, Inventory, and HR) into cloud data warehouses for reporting and analytics.  Built automated pipelines using SAP connectors, SQL, and Python, ensuring data integrity and reducing manual extraction efforts by 60% - Created and maintained dbt projects that provide clean data sets in Data bricks for enterprise-wide consumption  - Followed CICD best practices, including reviewing & critiquing Github pull requests to maintain a consistent, high-quality code base - Created and maintained extensive documentation on data sources, data models, and developer best practices/coding style guides  - Collaborated with subject matter experts to model data in a way that empowers analysts to answer their own questions. | | | | |
|  | Technical Advisor & Team Lead, Data Services Federal Government of Canada | | | | 2013-2019 |  |
|  | Designed and deployed a centralized data lake solution for multiple government departments, enabling cloud storage and advanced analytics.Developed ETL workflows using Apache Airflow and Spark, processing large-scale batch and real-time data streams. Mentored data engineering teams on best practices for data ingestion, processing, and visualization.Implemented agile methodologies to improve project delivery timelines and team collaboration. Managed/Administered Hadoop clusters using Cloudera Manager  Implemented security on Hadoop clusters (Kerberos, Knox, Ranger and Integration with Active Directory)  Developed ETL workflows using Oozie  - Built and deployed Azure Data Factory pipelines to move large volumes of data from on-prem locations to Azure Blob & Azure Databricks on a schedule - Enabled custom manipulation of data files in Azure Data Factory using Azure Batch - Optimize SQL query logic to improve execution times for reports that pull from on-prem and cloud-based databases - Collaborate with database developers, data architects, and QA engineers to promote data integrity and advance business intelligence  - Developed Python and SQL scripts to automate repeatable tasks, which saved the team ~20 hours per month - Designed public-facing cartographic products using ArcGIS, including a “Street Tree Density” map | | | | |  |
|  | Senior IT Customer Support Manager Kellogg, Brown and Root KBR Inc | | 2010-2013 | | |  | |
|  | ERP Migration Strategy & Planning, Develop and lead the end-to-end strategy for migrating from SAP to Oracle EBS, ensuring seamless business continuity during the transitionCollaborate with business unit stakeholders across HR (Hire to Retire), to understand legacy system requirements and ensure alignment with Oracle EBS functionalitiesCoordinate cross-functional teams (business analysts, developers, data engineers) to ensure all business processes are migrated and optimized within Oracle EBS Work with executive leadership and key stakeholders to define KPIs that measure the success of the ERP transformation | | | | |
|  |
|  |  | Web & DataBase Specialist UAE University (Ministry of Education) | | 2007-2010 | | |
|  |  | Working as a software engineer ensuring ERP, CRM and all web applications are designed, developed using highest standards | | | | |
|  |  | Senior Radio Frequency Planner Nokia Network. (https://www.nokia.com) | 2005-2007 | | | |
|  |  | Working as a Radio Frequency planner to implement telecom network across a region. Automating and programming telecom towers and building data pipelines for reporting purpose | | | | |
|  |  |  | | | | |

# EDUCATION AND TRAINING

**Master of Engineering Degree (Engineering Systems & Computing**)

University of Guelph (ON)  
Artificial Intelligence, Engineering Leadership, Robotic System, Machine Vision, Machine Learning

**Master of Information Technology**

University of Turku (Finland)

Project Management, Data mining, Cyber Security, Advance Internet of things, Information Security

**Bachelor of Computer Engineering**

NED University of Engineering and Technology

Software Engineering, Networking, Databases, Systems Engineering, Communication Systems  
 **Awards & Recognitions**

**Sheikh Hamdan award by Abu Dhabi, Gov. of UAE 2007**Innovation in Healthcare Award the Innovation in Healthcare Award is presented to projects that demonstrate outstanding innovation in the field of healthcare.

**Deputy Minister – Award of Excellence by Federal Gov. of Canada 2021**Excellence in Innovation, recognizes individuals who have accomplished outstanding & significant contribution to the furtherance of Government of Canada.

Other cetifications

* Six Sigma – Green Belt, Yellow Belt
* Cloudera – CCA 159 – Data Analyst using Sqoop, Hive, Impala
* Cloudera – CCA 131 – Cloudera Certified Hadoop and Spark Administrator
* Oracle Big Data Appliance - Administration
* .NET Framework 2.0-Web-based Client Development (**70-528 - MCTS**)
* .NET Framework 2.0-Application Development Foundation (**70-536 - MCP**)
* PRO: Designing and Developing Web Applications by Using the Microsoft .NET Framework (**70-547 – MCPD),** ITIL Foundation v3, PMP (Registered)
* PMP Certified (PMI ID: 6307991)
* Certified Ethical Hacker (2013)
* CCISP (2014)