**Rakesh Kumar Reddy Poreddy**

**+1(216)9033355 | LinkedIn:[LinkedIn/In/Rakesh](https://www.linkedin.com/in/rakeshkumarreddy-poreddy/) | [rakesh22rreddy@gmail.com](mailto:rakesh22rreddy@gmail.com)**

**Rakesh Kumar Reddy Poreddy**

**+1(216)9033355 | LinkedIn:[LinkedIn/In/Rakesh](https://www.linkedin.com/in/rakeshkumar-reddy-poreddy) | [rakesh22rreddy@gmail.com](mailto:rakesh22rreddy@gmail.com)**

**PROFILE SUMMARY**

A seasoned **Data Engineer** with over 10+ years of rich, hands-on experience in designing, building, and optimizing large-scale data solutions, Proven track record in designing, deploying, and managing secure identity platforms, with a specialization in **Microsoft Entra ID (formerly Azure Active Directory)**. Skilled in integrating identity solutions with complex enterprise systems, ensuring compliance, security, and scalability. to create high-performing, scalable, and cost-effective data solutions that empower businesses to make informed decisions.

I bring a versatile technical skill set, with expertise in **Big Data frameworks** like Apache Spark, Snowflake, Databricks, and Hive. My experience extends to workflow orchestration tools such as Apache Airflow and Informatica, which I use to streamline and automate data processes. With strong programming skills in **Python, Java, Scala, and SQL**,**SSIS.**I develop efficient and automated solutions tailored to complex business challenges. Additionally, I am well-versed in **DevOps practices**, utilizing tools like Terraform, Docker, and Kubernetes to ensure seamless **CI/CD**, infrastructure automation, and deployment consistency.

Throughout my career, I’ve successfully designed and implemented scalable data pipelines that handle multi-terabyte datasets, optimized both real-time and batch processing frameworks, and led critical migrations from legacy systems to cloud-based infrastructures with zero data loss. My approach to data engineering emphasizes data quality, governance, and compliance, ensuring adherence to industry standards such as **GDPR** and **PCI DSS**.

Collaboration is at the heart of my work. I have mentored teams, introduced best practices for cloud technologies, and fostered a culture of innovation and learning, enabling organizations to stay ahead of technological trends. Whether it’s creating robust data governance frameworks with **Databricks Unity Catalog**, designing secure multi-tenant storage solutions in **AWS S3**, or implementing cost-saving strategies, I take pride in delivering impactful, sustainable results.

I’m equally passionate about making data accessible and actionable for stakeholders. My experience with data visualization tools like **Tableau,Power BI**, and **QlikView** has helped bridge the gap between technical complexity and business insight, enabling leadership teams to make informed, strategic decisions.

To further validate my expertise, I hold several certifications, including **GCP Professional Data Engineer**, **Snowflake Certified Architect**,**AWS Certified Solutions Architect**, and **Databricks Certified Data Engineer Associate**. These credentials reflect my commitment to staying at the forefront of the ever-evolving data engineering landscape.

Driven by a passion for solving complex problems and enabling businesses to unlock the value of their data, I continue to focus on delivering innovative, high-impact solutions that empower organizations to thrive in a data-driven world.

**CORE COMPETENCIES**

* **Cloud Platforms:**
* GCP: **BigQuery, Dataproc, Cloud Dataflow, Cloud Pub/Sub, Cloud Composer**.
* AWS: **Redshift, Glue, DynamoDB**.
* Azure: **Data Lake, Cosmos DB**.
* BigDataFrameworks:**ApacheSpark,Snowflake,Databricks,ApacheKafka,Apache,Hive,Sqoop,Oozie**
* WorkflowOrchestration&ETL/ELT:**ApacheAirflow,Alteryx,DBT,AWSGlue,Talend,ApacheNiFi,Pentaho**

**Informatica**.

* Programming Languages: **Python, Java, Scala, SQL,Shell Scripting,SSIS.**
* Scripting Languages: **Shell, Perl.**
* Database Management: **MySQL, PostgreSQL, Teradata, MongoDB, Azure Cosmos DB**.
* DevOps & CI/CD: **Terraform, Docker, Gitflow, Jenkins, Bamboo**.
* Data Visualization: **Power BI, Tableau, QlikView, AWS QuickSight,**
* Agile Methodologies: **Scrum, JIRA, Confluence**.
* Certifications: **GCP Professional Data Engineer, Snowflake Certified Architect, AWS Certified Solutions Architect, Databricks Certified Data Engineer Associate**.

**SKILLS**

|  |  |
| --- | --- |
| * **Cloud platform serivces** * **Big Data and Data Engineering** * **Databases** * **Programming and Development** * **DevOps and CI/CD** * **Visualization and Reporting** * **Machine Learning and AI** * **Project Management and Collaboration** | Azure (Entra ID, Azure AD B2C, Azure AD DS, Azure DevOps), AWS (IAM, Cognito)Microsoft Entra ID, Azure AD B2C, Azure DevOps.  Apache Spark,PySpark,Apache Kafka, Hadoop,Apache Hive,Apache NiFi  MongoDB,AWS DynamoDB,PostgreSQL,SQL Server  Python,SQL Java Spring Boot  Terraform,Docker,Kubernetes,Azure Login Apps.  Tableau,Power BI  PyTorch  JIRA |

**PROFESSIONAL EXPERIENCE**

**Client: German Town Technologies | Portsmouth, VA**

**Sr. Entra ID Architect April 2023 – Present**

* Designed and deployed **Microsoft Entra ID solutions**, enabling secure identity access across hybrid environments.
* Implemented Conditional Access and **MFA policies**, improving organizational security posture.
* Migrated legacy identity management systems to **Entra ID**, ensuring seamless integration with enterprise applications.
* Established governance frameworks using Identity Protection and Privileged Identity Management (PIM).
* Designed and implemented scalable data pipelines using **Azure DataFactory** and **Databricks** for efficient ingestion, transformation, and storage of critical business data.
* **Migrated** legacy data assets to **Unity Catalog,Alteryx** establishing centralized data governance, lineage tracking, and secure access controls.
* Optimized **Databricks clusters** for cost-effective processing of large datasets, improving performance by **25%.**
* Developed role-based access control **(RBAC)** policies and implemented data encryption practices to secure sensitive data in Azure.
* Designed data models to support business intelligence needs in **Azure Synapse**, enabling actionable insights.
* Automated workflows using **Apache Airflow** to streamline ETL processes, reducing manual intervention by 30%.
* Led Agile sprint activities and mentored junior engineers on **Azure** and **Databricks** best practices, fostering a culture of collaboration and innovation.
* Designed and deployed monitoring tools and **CI/CD pipelines** to maintain consistency, scalability, and efficiency across cloud environments.
* Built a collaborative environment, regularly conducting training sessions, providing technical guidance, and enabling the team to stay ahead of evolving cloud technologies and data engineering trends.

**Environment:**

**AWS(EMR,Glue,S3),Snowflake,UnityCatlog,GCP,Python,SQL,SSIS,Databricks,ApacheKafka,,Airflow,Mongo,PostgreSQL,Alteryx,Teradata,DBT,Pandas.**

**Client: QBE | New York, NY**

**Data Engineer January 2021 – March 2023**

* Utilized **Docker** and **Kubernetes** to develop reproducible environments for consistent data pipeline deployments. Automated AWS provisioning with Terraform, enabling scalable and cost-effective solutions.
* Architected and administered **Azure AD B2C** environments, developing custom policies and enabling secure integration with enterprise applications.
* Automated identity lifecycle management using **PowerShell** and **Azure Logic Apps**.
* Collaborated with stakeholders to understand and fulfill identity management needs.
* Conducted regular system audits and implemented security enhancements.
* Designed and deployed secure **multi-tenant storage** solutions in **Azure Data Lake**, ensuring governance and **cost efficiency**.
* Streamlined data consistency by implementing **Terraform,SSIS** and **Python**-based automation for data transformations, reducing manual effort.
* Improved reporting accuracy in financial services by integrating **Informatica** and **Apache NiFi** for data quality management.
* Deployed real-time streaming solutions with **Databricks** and **Kafka** to ensure immediate data availability for business-critical decision-making.

**Environment:**

**AWS Glue,S3,Dynamo,QuickSight,DB,GCP,Databricks,Unity Catalog,BigQuery,GCS, Apache Spark,Shell,Terraform,Alteryx,Python,Informatica,Apache NiFi,PyTorch,JIRA,SQL,Git, Apache Kafka,Confluence.**

**Client: PNC Bank | Pittsburgh, PA**

**Data Engineer October 2018 – December 2020**

* Designed, built, and maintained robust **ETL (Extract, Transform, Load)** pipelines to process high-volume financial data, ensuring efficient data flow and availability for analytics teams.
* Designed and implemented robust **ETL pipelines** in **Azure Synapse** to support high-volume financial **data processing**.
* Migrated Hadoop workflows to **Azure Databricks**, achieving significant improvements in scalability and performance.
* Developed role-based **data security frameworks** and encryption strategies for sensitive financial datasets.
* Created data models and dashboards in **Tableau and Power BI**, enabling data-driven decision-making for executive teams.
* Implemented data quality checks and governance policies, ensuring compliance with regulatory standards like **GDPR** and **PCI DSS**.
* Proficiently worked with **Python**,**SQL,SSIS,Spark**,**Hadoop** and **Tableau** for data engineering and visualization tasks.

**Environment:**

**Apache(NiFi,Kafka),AWS(Glue,S3),DynamoDB,GCPDataproc,Terraform,Jenkins,Python,Docker,Git,**

**Informatica,QuickSight,JIRA,PyTorch,Tableau.**

**Client: Exl Service.com (I) Pvt. Ltd. | India**

**Data Quality Analyst April 2017 – June 2018**

* Analyzed data quality issues and implemented improvements using **Python, SQL**, and **Apache Hive** for comprehensive **big data** insights.
* Designed and developed dashboards and reports in **QlikView** and **Tableau**, enabling data-driven strategic decisions. Managed collaborative development of data projects using **Subversion** (SVN).
* Enhanced reporting capabilities by building robust data models and interactive **Power BI** dashboards. Utilized **Apache Hadoop** to process large datasets, improving data access and analysis efficiency.
* Conducted quality assurance for large-scale datasets processed on **Apache Hive** and **Hadoop.**
* Strong experience in using **Pentaho** **Data Integration** (PDI) to develop and implement ETL processes for large-scale data integration projects.
* Developed interactive reports in **Tableau** and **QlikView**, enabling data-driven decision-making.
* Automated data validation processes using **Python** and **Shell** scripting, reducing manual effort by 40%.
* Tuned **SQL** queries to enhance data retrieval speed and response times. Managed data migrations and integrations using **Apache Hadoop**, maintaining high levels of accuracy.
* Developed **SQL** scripts to streamline database management and reporting activities. Delivered detailed data analysis reports to stakeholders, influencing business strategies.

**Environment:**

**Python,PDI,SQL,Shell,ApacheHive,QlikView,Tableau,Subversion(SVN),PowerBI,ApacheHadoop,JIRA**.

**HighRadius Technologies, India**

**SQL Developer July 2015 – March 2017**

* Designed, developed, and managed **SQL** databases using **MySQL** and **PostgreSQL** to support business analytics and reporting. Configured **PostgreSQL** databases to enhance performance, scalability, and security.
* Utilized **Talend** to implement efficient **ETL** workflows, integrating data from various sources to enable robust business intelligence solutions.
* Led seamless data migration projects, ensuring minimal downtime and data integrity.
* Analyzed and optimized **SQL** queries and database performance to improve processing speed and reduce resource consumption.
* Developed and maintained **Power BI** dashboards, delivering actionable insights to support decision-making processes. Wrote complex **SQL** queries to facilitate in-depth data analysis and generate detailed reports.
* Partnered with business analysts to understand data requirements and provide tailored database solutions. Trained and mentored junior developers in **SQL** programming, database management, and **ETL** best practices.
* Managed database backups using Git for version control, ensuring data availability and integrity. Documented database designs, **ETL** processes, and data management workflows to establish a knowledge base for future use.

**Environment:**

**MySQL, PostgreSQL, Talend, Git, Power BI, Python**.

**EDUCATION**

* **Guru Nanak Institution Technical Campus**