📞(+1) 330-554-5757 📍 KENT, OH | ✉️koppula199@gmail.com | 🔗 linkedin.com/in/nagu-reddy

# Nagendra Reddy Koppula

*Senior Data Engineer*

## PROFESSIONAL SUMMARY

Overall, 5 years of experience as a Senior Data Engineer skilled in architecting and implementing scalable data solutions across AWS, Azure, and GCP. Proficient in integrating AI/ML models Vertex AI, Azure ML into data pipelines to generate business insights. Experience includes building end-to-end ETL workflows, data lakes, and streaming analytics for diverse industries healthcare, telecom, automotive, insurance. Good Communication, Presentation, Decision Making Skills, Understanding Business Requirements, Good Team player and always tries to lead the team with good efficient output.

## PRIMARY SKILLS

***ETL/ELT Processes, Azure Data Factory, Azure Databricks, Azure Data Lakes, Azure Data Tables, Python, Azure Synapse Analytics, Azure Blob Storage, Azure SQL Database, Azure Machine Learning, Data Warehousing, Data Modeling, Big Data Processing (Spark, Hadoop), Google Cloud (BigQuery, Dataflow, Pub/Sub, Composer), AWS (Glue, Redshift, Lambda, S3, EMR).***

## EMPLOYEMENT HISTORY

### SENIOR DATA ENGINEER

*Startech Networks Inc, Plano, TX, USA* *Sep 2025 – Present*

- Led the design and implementation of end-to-end Big Data pipelines 25% using Cloud Dataflow, Dataproc (PySpark/Spark), BigQuery, and Cloud Storage, ingesting structured and unstructured data to support analytics and AI workloads.
- Developed scalable ETL and transformation frameworks 20% with Apache Beam, Python, SQL, and Dataproc, ensuring incremental loads, data quality, and schema evolution for multi-terabyte datasets.
- Architected and optimized BigQuery data warehouses 15%, implementing partitioning, clustering, materialized views, and performance tuning, reducing query/reporting times by 40%.
- Collaborated with data scientists and UX researchers (10%) to productionize ML models using Vertex AI, integrating machine learning outputs into reporting dashboards and automated pipelines.
- Managed client-facing technical engagements 10%, providing consulting on GCP architecture, Big Data solutions, and pipeline optimization, while troubleshooting complex data issues.
- Implemented CI/CD automation and version control 10% using Cloud Build, Terraform, and GitHub Actions, deploying Dataflow jobs, Dataproc notebooks, and BigQuery scripts efficiently and reliably.
- Mentored junior engineers and promoted best practices 10% in ETL/ELT design, data governance, data lakehouse architecture, AI/ML integration, pipeline monitoring, and performance optimization, enhancing team productivity and delivery quality.

### AZURE DATA ENGINEER

*Subaru of America, Delaware, USA.* *July 2024 – Aug 2025*

- Developed and optimized data pipelines 25% using Azure Data Factory and Databricks (PySpark), improving large-scale data ingestion, transformation, and ML feature preparation by 35%.
- Built and automated ETL/ELT workflows 20% integrating Oracle, Dynamics 365 CRM, and telematics data into Azure Data Lake and SQL Database, increasing data availability for analytics and ML training by 40%.
- Implemented Lakehouse architecture 15% in Databricks with Delta Tables (Bronze, Silver, Gold layers), reducing query times by 30% and supporting AI/ML model pipelines.
- Applied Unity Catalog 10% for unified governance, metadata management, and access control, improving data consistency and compliance for analytics and ML datasets by 25%.
- Leveraged GitHub for CI/CD and version control 10%, streamlining code and ML model deployments, boosting release efficiency by 20%.
- Automated Azure infrastructure setup 10% with Terraform and ARM templates, ensuring scalable and repeatable environments for data pipelines and ML workflows with 30% faster deployments.
- Collaborated with Data Science and EDW teams to optimize data models and ML features, enhancing predictive analytics accuracy and ML model reliability by 20%.

### DATA ANALYST

*Kent State University, Kent, OH, USA.* *Feb 2023 – May 2024*

- Developed and automated interactive Power BI dashboards that improved data visibility and reporting efficiency by 45% across multiple university departments.
- Analyzed student performance and enrollment data to identify key academic trends, supporting data-driven decisions that increased retention rates by 12%.

- Optimized data extraction and transformation processes using SQL and Power Query, reducing data processing time by 35%**.**
- Collaborated with academic and administrative teams to design KPI-based visualizations, improving operational transparency and decision-making accuracy by 30%**.**
- Integrated Power BI with Excel and SharePoint to streamline report distribution, cutting manual reporting efforts by 40%**.**
- Improved data accuracy and consistency by 25% through validation checks, ETL automation, and establishing standardized data governance procedures.

### ASSOCIATE DATA ENGINEER

*BeyondScale - Hyderabad, India* *Nov 2020 – Dec 2022*

- Migrated on-premises insurance data warehouse to AWS, consolidating data into Amazon S3 and Redshift, which reduced reporting time by 50%.
- Designed serverless ETL pipelines with AWS Glue and Lambda for real-time claims processing, reducing average claim processing time from 4 hours to 15 minutes.
- Developed Redshift-based analytics platform for dynamic risk analysis and fraud detection, improving detection rates by 20%.
- Implemented AWS best practices (IAM roles, S3 encryption, VPC) for security and architecture, achieving full compliance with insurance data regulations.
- Automated data lake ingestion workflows using AWS Glue jobs and Lambda triggers, eliminating manual steps and reducing data latency by 80%.
- Collaborated with cross-functional teams to deploy infrastructure as code (CloudFormation/Terraform), accelerating environment provisioning by 3x.

## TECHNICAL PROFICIENCY

**Cloud Platforms:** Microsoft Azure, Azure Data Factory, Azure Synapse Analytics, Azure Databricks, Azure Data Lake, Azure Blob Storage, Azure SQL Database, Google Cloud (BigQuery, Dataflow, Pub/Sub, Composer), AWS (Glue, Redshift, Lambda, S3, EMR)
**Data Engineering:** ETL/ELT processes, Data Warehousing, Data Modeling, Big Data Processing (Spark, Hadoop), Amazon Redshift, Google Big Query.
**Programming Languages:** Python, PySpark, SQL Server, MySQL, PostgreSQL, MongoDB, Java, Shell Scripting
**AI/ML:** Azure Machine Learning, TensorFlow, PyTorch, , MLOps, NLP, GPT Models, Vertex AI, Scikit-learn
**Tools & Frameworks:** Apache Spark, Apache Kafka, Kinesis, Power BI, Tableau.
**DevOps:** Git, GitHub, GitLab, Bitbucket, JIRA, Azure DevOps, CI/CD Pipelines, Terraform, Jenkins, Ansible Kubernetes, Docker.
**Monitoring:** Stackdriver, ELK Stack, Azure Monitor, Cloud Logging
**Other Skills:** AWS IAM , Data Governance, Azure key Vault, Data Security, APIs, RESTful Services, Event-Driven Architecture.

## EDUCATION

Masters in **Computer Science** from Kent State University. May 2024
Bachelor's Degree in **Electronics & Communication Engineering** from BIHER, India. May 2022