

# Vishnu Nagesh Kumar

Boston, MA | (917)-353-7058 | vishnukumartalanki@gmail.com | [LinkedIn](#)

## SUMMARY

- Professional engineer in the data, AI/ML space with 8 years of experience in the industry. Certified scrum master skilled in managing large teams and collaborating with cross-functional collaboration.
- Proficient in Python, SQL, PySpark, Databricks, data modelling, data governance and Snowflake on Azure, AWS and GCP environments. Strongly focused on building end-to-end scalable ETL pipelines.
- Strong hands-on experience working with LLMs (OpenAI, LLaMA, Anthropic, etc.), retrieval augmented generation (RAGs) and vector databases and semantic search.

## SKILLS

- Programming : Python, SQL, PySpark, JavaScript, Bash, Go, HTML5, CSS, C#
- Data : Snowflake, Databricks, BigQuery, DataProc, Redshift, Azure Synapse, dbt, Fivetran, Airflow
- AWS : S3, EC2, Lambda, Bedrock, Polly, CDK, API Gateway, ElastiCache, SNS, RDS, EKS CloudWatch, Athena, LakeFormation, Step Functions, SQS, Kinesis, ECS
- Azure: Blob Storage, Virtual Machines (VMs), Functions, , API Management, Cache for Redis, Service Bus/Notification Hubs, Azure SQL , Azure Database, AKS, Azure Monitor, Event Hubs, ACI
- AI/ML: GPT3.5, GPT4o, LLaMA3.1, ElasticSearch, Pinecone, Ollama, OpenWebUI, LangChain, LlamaIndex, LangGraph, Deepseek-r1, browser-use, crawl4ai, prompt engineering (zero shot, few shot, chain of thought, etc.), Mflow
- Other : Docker, Terraform, Jenkins, Git, Tableau, FastAPI, Streamlit, Postman, Pydantic, web scraping, neo4j, GraphQL, GCP

## CERTIFICATIONS

- Microsoft: Azure AI Engineer Associate
- Scrum.org: Professional Scrum Master (PSM-II)
- Databricks: Generative AI Fundamentals
- Meta: Introduction to Databases
- IBM: Design Thinking Fundamentals

## EXPERIENCE

### Chewy, Boston, MA

Lead Data Engineer, Enterprise Data Lake | July 2024 – Present

- Minimized ETL costs for Postgres and Dynamo replication workflows from \$140k to \$90k monthly (20% reduction) by switching to a new medallion architecture.
- Customized terraform to provision ECS containers and S3 instances for LakeFormation of new medallion ETL workflows of bronze/silver/gold tables.
- Fostered cross-functional collaboration, working with analysts, product managers, and engineering teams to align business logic with data products.
- Coordinated JIRA story requirements, acceptance criteria and test plans with Enterprise Data Lake, Data Ops and offshore contractors to facilitate design and development.

### Fox Sports, New York, NY

Senior Data Engineer, Sports Data Platform | Apr 2023 – Apr 2024

- Implemented a sports data and ranking platform on AWS/Databricks Data Lakehouse using PySpark, Pydantic and Spark SQL, integrated 1 TB+ semi-structured data (10 mn+ records) from 4 vendor REST APIs for 5+ leagues (MLB, NFL, etc.); enabled access to enriched data via 5 GraphQL APIs, handled 10000+ requests/day for services.
- Developed generative AI applications integrating sports data with Amazon Bedrock (LLAMA2) and Polly, driving engagement in live broadcasts.
- Collaborated with ML engineers and analysts to design metadata management solutions for real-time data systems.
- Managed a team of 4 offshore data engineers/python developers, utilized a kanban approach to delegate and plan ELT tasks and achieved 90% of planned sprint goals.

### Deloitte, New York, NY

Data Engineer Consultant, Strategy & Analytics | Apr 2021 – Apr 2023

- Engineered an end-to-end ETL pipeline using Python, Hive, SQL and DataSwarm (proprietary tool) to pull org-wide developer survey unstructured data for 13000+ developers across 150+ teams from Qualtrics to Hive on a real-time basis.
- Devised an ELT pipeline to load semi-structured batch data onto Snowflake from AWS RDS, AWS DynamoDB, etc. Leveraged dbt to run procedures to hide PII data in accordance with CCPA and GDPR. Achieved compliance with over 30+ PII fields.
- Collaborated with product and engineering teams to implement scalable pipelines for real-time data ingestion and analysis.

### United Nations, New York, NY

NLP Engineer, Crisis Bureau | Sep 2019 – Sep 2020

- Supervised a team of 8 consultants to execute an ETL/NLP pipeline to find the correlation between mainstream news sentiment and political instability/insurgency through Python, GCP, Airflow, BigQuery and DataProc; scraped over 10 million+ news articles for 7 countries in the Sahel region over a 20-year period.
- Produced 3 dashboards in Tableau to capture terrorism events and news sentiment for Nigeria, Burkina Faso, and Mali; highlighted 65 major events with a >0.7 correlation.

### Great West Financial, Bangalore, IN

Data Engineer | May 2017 – Sep 2019

- Pioneered a redesign of Oracle forms into a cloud-based web application with AWS and react.js; refined user experience and lessened manual efforts by 5000+ hours/week.
- Improvised an ETL solution to automate Salesforce requests to JIRA stories by orchestrating AWS Lambda and DynamoDB; reduced manual efforts by 90%.

## EDUCATION

- Columbia University, New York, NY – M.S. in Management Science & Engineering | Dec 2020
- B.M.S. College of Engineering, Bangalore, IN – B.Eng. in Electronics and Instrumentation | May 2017
  - Awarded Best Project for developing a real-time text-to-braille converter using Python and OpenCV.