Asha Dasi

Enterprise Cloud Data Architect–Ethical Data Science/Engineering/Gen AI

614-992-8709

rohitb@interaslabs.com

#

# SUMMARY

* Productive and focussed data professional with 14 years of experience and 5 years of developing and implementing enterprise data architecture standards, optimizing database performance, and improving data accuracy. Passionate about leveraging technical expertise to architect robust data solutions that foster business growth and efficiency. Recognized for innovative projects that focus on **ethical AI/ML data practices** and upscale data governance for privacy,security and compliance.
* Skilled in various cutting-edge technologies such as Artificial Intelligence, Machine Learning, Deep Learning, Natural Language Processing(NLP), ChatGPT, LLM, Data Analysis, Predictive Analysis, Data Mining, Data Visualization, Business Intelligence, and AWS MLOps. Possesses comprehensive industry knowledge across Healthcare and Life Sciences, Entertainment, Retail, and Telecom sectors.
* Proven track record in reducing data processing time by 60%, increasing data accessibility, and improving data security compliance by 45%. Adept at collaborating with stakeholders, researching new technologies, and implementing data quality checks to drive successful product launches and improve data analysis capabilities.

# SKILLS

* **Data Modeling**: NomNoml,Microsoft Visio, PowerDesigner,ER/Studio, IBM Data Architect, Oracle SQL Developer, MySQL Workbench
* **Big Data & Cloud** : Spark,Hadoop, MapReduce, Hive, VertexAI, Airflow, Snowflake, BigQuery, Redshift, PostgreSQL, DynamoDB, Cassandra, MongoDB, MySQL, Oracle, AWS, Google Cloud(GCP), Azure ,AWS SageMaker
* **Programming & Scripting**: Pyspark,Python (Pandas, NumPy), Scala, SQL, Machine Learning, Deep Learning, NLP ,MLOps,PL/SQL, Unix
* **Data Analysis & BI Tools**: Tableau, Power BI, DAX Studio, Google DataStudio, Tableau & SAP BI, QlikView
* **Methodology & DevOps:** Agile (JIRA & Confluence), CI/CD, Git, Jenkins, Dockers, Kubernetes, Terraform

## OTHER

* **Certifications**: Pursuing Certified Ethical Emerging Technologist Professional Certificate(**CEET**) and Project Management Professional (**PMP**)
* **Technical Mastery:** Continuous education on latest tech-The Open Group architecture Framework(**TOGAF**), Gen AI,NLP,LLM
* **Professional Associations:**Volunteered for DataConLA and member since 2018,Learn to Code meetups/conferences in Pasadena and LA

# EDUCATION

* Bachelor of Technology in Electrical Engineering, JNTU 2007
* Master’s in Computer Science, Texas A&M University Commerce, TX 2009

# EXPERIENCE

**Davita(Remote) Dec 22 - Present**

# Enterprise Data Architect/Lead Data Consultant

* Engineered a real-time data sync framework between operational and analytical systems,amplifying reporting time by 70%.
* Harmonized disparate data sources from multiple acquisitions into a cohesive data lake, simplifying analysis and insight generation for the leadership team.
* Crafted and enacted a disaster recovery strategy that reduced potential data loss scenarios by 85%, safeguarding critical business intelligence related to clinical trials and measures of patients with ESKD.
* Build Airflow feature,prediction and implementation dags to train ML models and run them in hybrid mode - on-prem and AWS cloud.
* Directed the enterprise-wide modernization of data architecture aligning with digital transformation goals, projecting a 50% ROI over the next three years.
* Led a 10-member team to redesign the company's data architecture, increasing system performance by 45% while decreasing maintenance overhead.Applied Statistics along with end to end ML engineering including Classification, Clustering and Regression in detecting product anomalies
* Created multiple dimensions, perspectives, hierarchies, and dynamic measures (DAX) to analyze performances of each KPI.
* Design, document and test FHIR Implementation Guides (IG’s) and FHIR extensions that will enable external partners—medical centers and vendors--to structure data submissions in a standard, repeatable way.
* Worked with Data Scientists to create a POC plan for existing ML jobs to run and test the data using python/pyspark.

**HighmarkHealth(Remote) Jan 22 - Dec 22**

# Infrastructure Cloud Architect/Lead Data Engineer

* Orchestrated large-scale Cloud transitions that increased application performance by 80%.
* Spearheaded a team of 10 in enforcing cloud infrastructure strategies leading to better adaptability for unforeseen tech expansion.Developed a master data management(MDM) protocol in collaboration with 5 department leads, ensuring data integrity and boosting operational efficiency.
* Evaluated cloud application requirements, reducing infrastructure inefficiencies by 40%
* Created performance tuning methods on measures (DAX) for the reporting team using DAX Studio andSQL Profiler.
* Piloted advanced data analysis methods using predictive analytics, contributing to a 20% increase in targeted marketing campaigns’ success rate.
* Led a project to assign/update a unique ID for all the clinical sources(event-based) to help Highmark customers with inpatient/outpatient claims.
* Created a unique MDM lookup service framework to get the ID back from the system and update the Hive database using Pyspark on a daily basis and on- need for event based data.

**DocuSign,Seattle,CA(Remote) Jul 21 - Jan 22**

# Senior Data Consultant

* Pioneered an ETL pipeline automating the ingestion of 4TB of data daily, improving reporting accuracy for stakeholders.
* Designed a self-service BI tool with BQ, leading to a 40% increase in productivity among non-technical users.
* Authored a comprehensive data dictionary that standardized definitions across the organization, fostering better inter-departmental communication.
* Collaborated with cross-functional teams to introduce a data governance framework that improved data quality by 70% and ensured regulatory compliance.
* Orchestrated the transition to a microservices-driven architecture, improving system modularity and facilitating a 45% faster time-to-market for new features.
* Worked on performance tuning of Snow SQL to process/cure billions of docusign events.

**Lacuna Technologies - Full Time(Remote) Nov 20 - Jul 21**

# Senior Data Architect/Engineer

* Architected a scalable data warehousing solution that supports real-time analytics, resulting in a 40% increase in data retrieval efficiency for client-facing applications.
* Rolled out a company-wide privacy compliance strategy for data management, ensuring adherence to GDPR and CCPA laws, avoiding potential fines.
* Utilized SQL and Python scripting to analyze 500M+ geo spatial data points, uncovering actionable insights that increased user retention.
* Spearheaded the migration of 10 TB of critical business data to a cloud-based platform using AWS Redshift, which cut costs by 30% annually.
* Optimized data processing procedures by implementing Apache Spark, reducing batch processing times by 50% while handling 100 million+ data points daily.

**Warner Bros, Burbank CA Feb 18 - Aug 20**

# Senior Data Engineer

* Executed efficient installation and maintenance of cloud systems, boosting system reliability and performance by 40%
* Worked on data warehouse transition from AWS Redshift to Snowflake improving data management by 30%.
* Worked on a job automation of CCPA/GDPR requests implementation to purge any PII data, saving up to 20 hours per week.
* Worked with 200TB and more data volumes of user data and clicks to provide meaningful data to the data science team.
* Built a touchless, automated, schema-dynamic pipeline to process files with no schedule using AWS Lambda,Python.
* Acted as product lead for the Harry Potter related app on iOS/Android coordinating with teams in the UK,US and India and successfully reached the stakeholders and users expectations.
* Implemented the solution using AWS EMR with Apache Spark(python) to process seamlessly and automatically the daily user reports for WB Consumer Intelligence team.
* Processed event files using Kafka queueing /Spark streaming,done aggregations using the window functions, utilizing watermarking feature in Spark streaming.
* Designed a pipeline using Pyspark to filter and scrub user data from Oracle Bluekai received every hour loaded into Snowflake for the data science team to build and train their ML models.
* Wrote Apache Spark scripts using Python for data analysis and data matching for POC with – Data bricks and Qu bole.

**University Of California, Irvine, CA**

**Comtech Tel, Aliso Viejo, CA Oct 16 – Dec 17**

# Big Data Engineer/Consultant

* Supported the lead data architect in developing an enterprise data model that accelerated the time-to-insight for business users by 20%.
* Collaborated in the migration of legacy systems to cutting-edge database technology, which increased data storage efficiency by 30%.
* Assisted in the implementation of a big-data Hadoop cluster handling over 20TB of data, leading to a 50% increase in data processing speed.
* Involved in the design of a robust ETL framework that streamlined data integration from multiple sources, culminating in boost to overall data quality.

**Illumina, San Diego, CA Mar 16 – Oct 16**

# Data Engineer/Consultant

* Conducted comprehensive reviews and updates to existing database systems, improving operational efficiency by 15%.
* Partner closely with subject method experts and management to develop intelligent data initiatives. This involves assessing technology readiness, understanding the latest trends in the field, and developing strategies and roadmaps to coherently drive capability build-up.
* Partnering with the business in delivery of solutions and results through technology enhancements, implementations, and enhancements.

 **DaVita , Brea, CA Feb 14 – Jan 16**

 **ETL Informatica Consultant**

* Performed Impact analysis on the whole warehouse when major design changes are implemented to improve performance of Fact/Dim jobs due to changing data size.
* Wrote NZSQL scripts to load fact and dimension tables in Netezza using Aginity Workbench.
* Wrote HQL queries to get monthly and yearly data like count of Inpatient/Outpatient, Mode of Dialysis, Check vitals, Patient risk factor

**OfficeMax, Naperville, IL(Full Time)**

**CenturyLink,Monroe,LA**

**College Board, Reston, VA Feb 10 – Jan 14**

**ETL Production Support Engineer/Informatica consultant**

* Played a key role in closing many critical issues/bugs in the high profit/sale season for the company.
* Re-designed ETL jobs for performance issues as part of Production Support on-call rotation basis.
* Developed custom item jobs to provide special prices on the products to customers with bulk orders all around the year.