Seshu Kumar Yalamati

Sr. Data Engineer

Contact

Plano, TX - 75024
+1(972)-520-6716
sreeseshuy@gmail.com

Key Skills

Data Engineering
Customer Data Platform
Technical Architecture
Analytical Approach
Management
Quick adaptive and
learning

Technologies

Python, Azure Dev ops
Snowflake, Data Bricks,
Apache Hive, Caspian
Cloud Computing,
Matillion Machine learning,
Apache Spark, Power BI
Teradata, Oracle, Sybase,
Azure Synapse and SQL
Server
Erwin, Power Designer
and Visio
Windows, Unix

Education

GNIIT Computers, 2003
*National Institute of
Information Technology
University*
Bachelor's degree
Computers, 2001

Objective

With over 18 years of extensive Information Technology experience across diverse domains such as Retail, Banking & Financial Services, Consumer Goods, Healthcare, and Defence, I bring a wealth of knowledge to the table. Throughout my career, I have accumulated over 15 years of specialized expertise in Data Warehousing, encompassing roles such as Data Engineer, Data Modeling, Database Governance, Customer Data Platform Analyst, Cloud Computing, Teradata ETL, and hands-on utilities such as Azure ADF, DevOps, Azure Databricks and Snowflake. Additionally, I have knowledge of machine learning, AWS, and a range of databases including Oracle, Sybase, Azure Synapse, and SQL Server. I am also proficient in programming languages such as Python and PySpark.

My ability to maintain unwavering focus on achieving bottom-line results is complemented by my proficiency in formulating and implementing advanced technology and business solutions that cater to a wide range of needs. As an active team player, I effectively communicate and collaborate with colleagues, leveraging my interpersonal skills to foster productive working relationships. Despite working under intense pressure and tight deadlines, I consistently deliver business-critical projects on schedule and within budget.

Experience

December 2021 – Till Date

Sr. Data Engineer: Genpact
Projects Domain: Banking domains
Client: Morgan Stanley Bank

PROJECT PROFILE:

Project Title	Cloud Migration / Fraud Alert
Duration	12/07/2021 – Till date
Location	New York USA: Remote
Hardware	Windows 7
Operating Systems	Unix , Cloud

Osmania University

Technologies	AWS Glue, Kafka ,Snowflake, Matillion, SQL Server, Azure Synapse, Collibra, Apache hive, Airflow, Python, PySpark,Spark and Visio

PROJECT DESCRIPTION:
CLIENT DESCRIPTION:
Morgan Stanley is an American multinational investment management and financial services company headquartered at 1585 Broadway in Midtown Manhattan, New York City.

PROJECT DESCRIPTION:
Presently engaged in a Cloud Migration / Fraud Alert initiative at Morgan Stanley Bank, focusing on real-time payment transaction processing, fraud detection, and generating alerts. My responsibilities encompass diverse areas, including data architecture, data flow management, adherence to data protection principles, and actively participating in the data modeling of Fraud Domain objects. This involves implementing the models through SQL Server, as well as designing and executing ML models for fraud deduction, utilizing PySpark. Moreover, I oversee data quality, establish data lineage, and manage the migration of data across various stages, from varous source systems through AWS, and finally to Snowflake.

Fraud Alert: (Cloud Project: Payload Source → SQL Server → hive (Caspian))

The initiating of the payload is carried out via the payment gateway, followed by processing through the Kafka framework. Subsequently, the persistence service consume the payload from Kafka and enhances the payload to facilitate advanced fraud detection procedures. Simultaneously, the relevant fraud alert topic is activated in the Actimize system, while the corresponding payload is stored within the data model to support subsequent processing stages.

Fraud Analytics: Cloud Project: Source DB's → AWS S3 →Snowflake

The E*Trade fraud data migration involves transferring data from on-premises sources like Oracle, DB2, Sybase, and file-based systems to Snowflake, utilizing the AWS Glue framework through a metadata-driven process. This migration is orchestrated by Control-M jobs, which trigger an AWS Lambda function workflow. The process begins by extracting metadata from a staging S3 bucket. Subsequently, corresponding Parquet files are generated in a Raw Zone S3 bucket, The data is then loaded into Snowflake Data Warehouse tables through AWS Glue jobs.

Roles And Responsibilities:
❖ Data Architecture Fraud Data flow which persists real time payment transactions and streamlines necessary data to Actimize case management system which ensure Data Governance, Data Lineage and Data Quality in every phase of process.
❖ Creating logical and physical data models based on business requirements and industry standards.
❖ Participated in the implementation of a Slowly Changing Dimension(SCD) Type 2 framework within Databricks, leveraging Delta Lake tables to guarantee data consistency, optimize performance, and uphold data quality standards.

- Integrated Airflow with data storage solutions like Apache Hadoop, Amazon S3
- Ensuring compliance with data governance policies and data quality standards in the data modeling process.
- Analyzing business requirements and data sources to understand data needs and identify relevant data entities.
- Contributed to the analysis of the data lake and the creation of an analytics layer, empowering data scientists to develop fraud detection models from the source Big Data data lake using Spark, PySpark, and Python.
- Leveraged Pandas and NumPy to perform comprehensive data cleaning, manipulation, and analysis on large datasets, resulting in more accurate and actionable insights.
- Designing and implementing data warehouse fraud solutions on cloud platforms Snowflake to support business intelligence and analytics needs.
- Integrating data from multiple sources into a cohesive and structured model.
- The Data flow from source to target will follow the data protection principles like data transparency, accuracy, storage, accountability, integrity and confidentiality
- Collaborated on the orchestration of data workflows and the integration of data from various sources into a cloud data warehouse, utilizing Matillion.
- Involve in implementing data integration and transforming complex fraud data using Azure Synapse Data Factory (ADF) for Data scientist analysis.
- Involved in developing testing SQL server data objects framework by python FAST API which will connect to SQL server pass test JSON inputs and generate respective reconciliation reports
- Involved in fraud deducting ML modeling design, PySpark involves in fraud enrich logic
- Profiling legacy data / real time data by Collibra OWL tool and alert anomalies in data processing to respective upstream teams
- Data Quality and data lineage is achieved in data migration from Source system → Oracle / Sybase → AWS → till Snowflake to achieve quality data to fraud analytics

September 2012 – December 2021

Technology Specialist. Cognizant Technology Solutions
Projects Domain: Retail and Banking domains
Client: PepsiCo, AMEX Bank, Wells Fargo Bank and J P Morgan Bank

PROJECT PROFILE:

Project Title	TPO / AP Reporting / HRDM
Duration	09/10/2018 – 12/06/2021
Location	Plano TX USA: Hybrid
Hardware	Windows 7
Operating Systems	Unix , Cloud

Technologies	Azure Data Brick, ADF, Synapse, Snowflake, Teradata, oracle, unix, python and Visio

PROJECT DESCRIPTION:

Worked on various projects, including Total Portfolio Optimization (TPO), Global HR Analytics and Reporting Platform (HRDM), and AP Reporting, for clients such as PepsiCo, AMEX Bank, Wells Fargo Bank, and J.P. Morgan Bank. These projects involved technologies like Azure Data Brick, ADF, Synapse, Snowflake, Teradata, Oracle, Unix, Python, and Visio. I played a key role in data analysis, data modeling, developing metrics, KPIs, analytics, and dashboards, as well as collaborating cross-functionally to analyze growth targets and inform priorities. I also contributed to customer data integration, ETL mechanisms and strategies using Teradata Utilities and Informatica Power Center Tool, and overall project delivery.

Roles And Responsibilities:

❖ Analyze and interpret complex data on systems, identify the Key Business Performance Indicator metrics and various key measures & dimensions and provide resolutions to data issues and coordinate with business analyst to validate the business requirements.
❖ The Data flow from source to target will follow the data protection principles like data transparency, accuracy, storage, accountability ,integrity and confidentiality
❖ Designing and implementing data warehouse solutions on cloud platforms Azure Synapse Analytics to support business intelligence and analytics needs.
❖ Developed interactive data objects layer for supply chain stakeholders, providing actionable insights for decision-making.
❖ Led data governance initiatives to ensure compliance with data privacy regulations and standardization of supply chain data across the data models.
❖ Lead the team on customized CDP tool which focus on identifying opportunities to improve customer loyalty program and track treasure data of hot products
❖ Data modeling of Data Acquire Layer, DWL Layer, App Layer and Semantic Layer
❖ Further develop metrics, KPIs, analytics and dashboards to provide marketing measurement and drive marketing performance through deep business insight
❖ Profiling data belong to customer buyer behavior, path to purchase or decision science models
❖ Develop and direct a variety of fact-based analyses that aid Management in making decisions relating to the creation or expansion of new or existing businesses
❖ Collaborate cross-functionally (Marketing, Finance, Stores, IM) to analyze on growth targets and inform priorities to CDP analytics team
❖ Improve the effectiveness of customer acquisition, customer wallet share deepening, customer value management and customer retention
❖ Manage multiple data sources, identify gaps in the data capture process and make suggestions on improvements/enhancements for data hygiene
❖ Provide data analytics to be used by the product management and sales team to make decisions regarding product strategy

- ❖ Responsible to develop transformations in Databricks and integrate into Azure ADF pipelines framework for customer data integration for data analytics
- ❖ Analyze on the customer level data and involve in data collection, data governance, data processing and data activation
- ❖ Design Develop and Architect the Extraction, Transformation and Loading (ETL) mechanisms and strategies using Teradata Utilities, Informatica Power Center Tool and Linux Shell scripts as per business requirements. Follow Architecture framework and best practices to design the ETL solutions to improve the data load/retrieval processes to meet business Service Level Agreements.
- ❖ Responsible for performing analysis for various enhancements and perform impact analysis to find out the systems/programs that could potentially be affected by the proposed changes. Coordinate with multiple teams involving the IT stake holders, business users and the technical onshore team for system testing, User Acceptance Testing and production issues resolution.
- ❖ Validate data flow job performance metrics, fine tune underperforming jobs to fall within approved operational performance limits (CPU usage, I/O operations etc.). Work with the Database Administrators to implement compression strategies for efficient data storage
- ❖ Unit and Regression testing of all developed artifacts where development defects are reported, tracked, fixed and retested, until the application reaches the successful completion of the User Acceptance testing.

Project Title	*Control KPI : (Cloud : Data Lake → Azure)*
Duration	*08/01/2016 – 09/09/2018*
Location	*Hyderabad*
Hardware	Windows 7
Operating Systems	Unix , Cloud
Technologies	Azure Data Brick, ADF, Synapse, Teradata, oracle, unix, python and Visio

PROJECT DESCRIPTION:
PepsiCo Inc. is an American multinational food and beverage corporation belong to Retail &Consumer Goods domain. The Control Key Performance Indicator (KPI) Project Establish an automated analytics model for monitoring of key risk and controls across 16 entities of PepsiCo AMENA and EUROPE Sector for procure to pay (P2P) cycle.

ROLES AND RESPONSIBILITIES:
Automation of the key controls monitoring activity thereby saving time on manual analysis and testing. Involved in Data Analytics and exceptions reporting finalization with the business, Develop technical blueprint of the operating model and identification of data source and mapping with the KPIs which includes acquire and integrate the data from source ERP

system to EDW system and Involved in innovation activities and got appreciations from customer
DURATION OF THE PROJECT: 3 + years

Project Title	CLT (On Perm informatica → Teradata)
Duration	01/10/2013 – 06/31/2016
Location	Hyderabad
Hardware	Windows 7
Operating Systems	Unix , Cloud
Technologies	Teradata, oracle, Informatica, unix, python and Visio

PROJECT DESCRIPTION:
Consumer Lending Technology Line of business deals with the projects where the data flow from operational systems to analytical systems of Mortgage loans. Involved in DDM, PPDW, FSR and AIS projects includes analyzing Source to target flow, designing the mapping documents and building conceptual and logical designs
CLIENT DESCRIPTION:
Wells Fargo & Company is an American multinational financial services company headquartered in San Francisco, California, with central offices throughout the United States.[10] It is the world's fourth-largest bank by market capitalization and the fourth largest bank in the US by total assets.

ROLES AND RESPONSIBILITIES:

❖ Involved in creating source to target mapping documents and Integration and semantic Validation. Handling data warehouse Architecture encompassing data marts design, validating origination data and preparing detail & summary reports and participating in business analysis discussions and overseeing overall project delivery. Involved in innovation activities and got appreciations from customer. The Data flow from source to target will follow the data protection principles like data transparency, accuracy, storage, accountability, integrity and confidentiality

DURATION OF THE PROJECT: 2.8 years

Project Title	Global Credit Administration (GCA) (On Perm informatica → Teradata)
Duration	01/10/2012 – 31/08/2013
Location	Hyderabad
Hardware	Windows 7
Operating Systems	Unix
Technologies	Teradata, oracle, Informatica, unix, python and Visio

PROJECT DESCRIPTION:

Outside Agency project deals with the defaulter's credit cards customers. The OA tools are split between Exception and Standard process. The defaulters are classified in to various tools like SCRA, AC and BK etc. The ultimate benefit is to analyze defaulter's credit cards customers by these OA tools.

CLIENT DESCRIPTION:

The American Express Company, also known as Amex, is an American multinational financial services corporation headquartered at 200 Vesey Street in New York City. The company was founded in 1850 and is one of the 30 components of the Dow Jones Industrial Average. The company is best known for its charge card, credit card, and traveler's cheque businesses.

ROLES AND RESPONSIBILITIES:

❖ Involved in Data modeling activities which included Source to Staging and Target. Developed ETL process using SSIS and data warehouse design phase and migrated SAS code to Teradata code and organized the data from actual IDN Tera Data Tables to Staging Layer Lead team activities and gave walk through to client in every SDLC phase. The Data flow from source to target will follow the data protection principles like data transparency, accuracy, storage, accountability ,integrity and confidentiality

DURATION OF THE PROJECT: 1 years

April 2010 – August 2012

Technology Lead. iGATE global solutions limited
Projects Domain: Banking
Client: RBC Bank

Project Title	Royal Bank of Canada (RBC) (On Perm Mainframe DB file → Teradata)
Duration	01/05/2010 – 31/08/2012
Location	Hyderabad
Hardware	Windows 7
Operating Systems	Unix
Technologies	Teradata, oracle and Mainframe

CLIENT DESCRIPTION:

The Royal Bank of Canada (RBC; French: Banque royale du Canada) is a Canadian multinational financial services company and the largest bank in Canada by market capitalization. The bank serves over 16 million clients and has 86,000+ employees worldwide.

ROLES AND RESPONSIBILITIES:

- ❖ Involved in analyzing scores of various models and preparing Tech spec Documents, Unit Test Plan Documents and Coding in Tera Data. Involved in performance tuning and using fastload. Involved in presenting walkthrough to client in every phase of SDLC. Development of Teradata query scripts and execute through JCL programs. Authentication testing of production results by executing JCL Jobs. Understanding the design and scope of the module.
- ❖ Profiling value of acquired customers by promotion type
- ❖ Profiling of customers responding to emails
DURATION OF THE PROJECT: 2.3 years

February 2009 – April 2010

Software Engineer. Prokarma
Projects Domain: Health and Real Estate domains
Client: Televox

Project Title	Televox Dotnetnuke
Duration	01/02/2009 – 11/05/2010
Location	Hyderabad
Hardware	Windows 7
Operating Systems	Windows
Technologies	Dot net nuke 4.9, Asp.net 3.5and jquery

ROLES AND RESPONSIBILITIES:
- ❖ Involved in coding using C#.Net, JQuery and Sql Server 2005. Understanding the design and scope of the module. Implemented Agile Software development like Scrum
DURATION OF THE PROJECT: 1.2 years

December 2007 – February 2009

Software Engineer. Power byte comp solutions
Projects Domain: Health and Real Estate domains
Client: Prokarma

Project Title	Prokarma
Duration	01/12/2007 – 31/01/2009
Location	Hyderabad
Hardware	Windows 7
Operating Systems	Windows
Technologies	Dot net nuke 4.9, Asp.net 3.5and jquery

- ❖ Developed the screens, wrote queries and accountable for coding and designing. Established installation and up gradation of packages using Install Shield 12
DURATION OF THE PROJECT: 2.1 years

March 2006 – December 2007

Sr. Software Engineer. Leela info tech Pvt. Ltd.
Projects Domain: Health
Client: Expanded Vision

Project Title	Lecture Reference Control
Duration	01/04/2006 – 28/02/2007
Location	Hyderabad
Hardware	Windows 7
Operating Systems	Windows
Technologies	C#.Net and ASP.Net

CLIENT DESCRIPTION:

Expanded Vision is an Designer Manufacturer Supplier since 68 Years as the Professional Pioneer Scientific Level Iridology Iridiagnosis Iridodiagnosis World Leader

ROLES AND RESPONSIBILITIES:

❖ Analyze and interpret complex data on systems, identify the Key Business Performance Indicator metrics and various key measures & dimensions and provide resolutions to data issues and coordinate with business analyst to validate the business requirements.
DURATION OF THE PROJECT: 1.1 years

March 2004 – March 2006

Programmer. Defence electronics research Laboratory
Projects Domain: Defence
Client: Defence Research & Development Organization

Project Title	Electronic Warfare
Duration	03/01/2004 – 03/31/2006
Location	Hyderabad
Hardware	Windows 7
Operating Systems	Windows
Technologies	C#.Net and ASP.Net

CLIENT DESCRIPTION:

Defence Electronics & Research Laboratory (DLRL) is a laboratory of the Defence Research & Development Organization (DRDO). Located in Hyderabad, it is actively involved in the design and development of integrated Electronic Warfare systems for the Indian Armed Forces.

ROLES AND RESPONSIBILITIES:

❖ Development of Teradata query scripts and execute through JCL programs. Authentication testing of results by executing JCL Jobs.
DURATION OF THE PROJECT: 2 years

Advance Technologies Exposed

❖ Azure Technologies

- Machine learning with ==Python== and Azure Notebooks
- Secure your Cloud ==Data==

❖ ==Data== Science
- ==Data== Research Practices

Seshu Kumar Yalamati

- Plano, TX, US

Contact Information

- tkb-t1n-mvv@mail.dice.com

- 1972520671

Work History

Total Work Experience: 20 years

- **Sr. Data Engineer Genpact**
  Dec 01, 2021

- **Sr. Data Engineer Cognizant Technology Solutions**
  Sep 01, 2012

- **Technology Lead Igate Global Solutions Limited**
  Apr 01, 2010

- **Software Engineer Prokarma**
  Feb 01, 2009

- **Software Engineer Power Byte Comp Solutions**
  Dec 01, 2007

- **Sr. Software Engineer Leela Info Tech Pvt. Ltd.**
  Mar 01, 2006

- **Programmer Defence Electronics Research Laboratory**
  Mar 01, 2004

Education

- **Bachelors** | Bsc Computers

- **Masters** | GNIIT

Skills

- **analytics** - 15 years
- **data profiling** - 15 years
- **data architecture** - 8 years
- **big data** - 7 years
- **teradata** - 13 years
- **data modeling** - 12 years
- **dimensional modeling** - 12 years
- **azure data factory** - 9 years
- **cloud computing** - 9 years
- **data collection** - 9 years
- **data engineering** - 9 years
- **data governance** - 9 years
- **data processing** - 9 years
- **data warehouse** - 9 years
- **etl architecture** - 9 years
- **leadership** - 9 years
- **retail** - 9 years
- **sales** - 9 years
- **transformation** - 9 years
- **utilities** - 9 years
- **business analysis** - 7 years
- **microsoft windows azure** - 5 years
- **cdp** - 3 years
- **python** - 4 years
- **banking** - 11 years
- **unix** - 9 years
- **cloud** - 2 years
- **oracle** - 2 years
- **it management** - 2 years
- **software engineering** - 4 years
- **dotnet nuke** - 3 years
- **jquery** - 3 years
- **programming** - 3 years
- **it** - 2 years
- **research** - 4 years
- **c#** - 3 years

Work Preferences

- Desired Work Settings: No Preference
- Likely to Switch: False
- Willing to Relocate: False
- Travel Preference: 0%
- Work Authorization:
    - US
- Work Documents:
    - Have H1 Visa
- Security Clearance: False
- Third Party: False
- Employment Type:
    - Full-time
    - Contract - Corp-to-Corp

Profile Sources

- Dice: https://www.dice.com/employer/talent/profile/7a1264f295406091f81c90c0df4f3870