**Overview of DRD**

During my internship, I gained insights into various AML solutions, specifically SAM (Suspicious Activity Monitoring), WLF (Watch List Filtering), and CDD (Customer Due Diligence).

These solutions cater to different Lines of Business (LOB), including correspondent banking, financing, gaming, insurance, retail banking, and retail broker-dealer sectors.

Each solution addresses specific regulatory requirements and operational challenges. SAM focuses on detecting suspicious activities in transactions, WLF screens customers against watch lists, and CDD ensures thorough due diligence and risk assessment.

2.Understanding the Structure of DRD:

Hierarchy of Priorities: The DRD prioritizes data requirements based on their criticality and impact on AML processes. High-priority elements are essential for compliance and risk mitigation, while lower-priority elements support additional insights.

Types of Entities: Entities such as Account, Customer, Transaction, and Branch are key components of the AML framework. Each entity has specific attributes that describe its characteristics and role.

Data Elements and Attributes: Each entity includes attributes defined by data type, permissible values, and constraints. For example, the Account entity has attributes like ACCOUNT\_ID, ACCOUNT\_TYPE, and ACCOUNT\_STATUS.

Reference Tables: These tables maintain consistency in attribute values. For example, the ACCOUNT\_CATEGORY table includes categories like ‘Savings’, ‘Checking’, and ‘Loan’, referenced by the ACCOUNT\_CATEGORY\_CD attribute.

Entity Relationships: Understanding how entities relate is crucial. For instance, an Account entity might be linked to a Customer entity, indicating multiple accounts per customer. Transactions can link to both accounts and customers, creating a comprehensive data network for AML monitoring.

**UDM – SAM key entities**

I focused on understanding the key entities and their relationships within the SAM (Suspicious Activity Monitoring)

**Entities and Relationships**: I learned about the various entities within the SAM system and how they interact with each other. This includes understanding the relationships between parties, accounts, transactions, branches, and other elements.

**Segmentation and Grouping**: Understanding how accounts and parties can be segmented and grouped based on different criteria helps in better monitoring and analysis.

**Hierarchical Structure**: Recognizing the hierarchical structure of data elements, from individual transactions to broader population groups, is crucial for effective AML monitoring.

This comprehensive understanding of the illustration helps me appreciate the complexity and importance of data relationships in AML systems. It also highlights the need for meticulous data management and segmentation to ensure effective monitoring and compliance.

**Data Pipeline**

End-to-End Data Flow: I gained a comprehensive understanding of how data flows from source systems through various processing stages to the final data repository.

Data Quality and Integrity: I learned the importance of validation, mapping, and cleaning processes in ensuring high data quality and integrity.

Efficient Data Handling: The bulk loading and staging processes are crucial for handling large volumes of data efficiently and preparing it for final storage.

Central Data Store (CDS): The CDS serves as the primary repository for fully processed data, ensuring it is ready for analysis and operational use.

This detailed understanding of the data pipeline highlights the complexity and importance of each stage in ensuring reliable and high-quality data management.

5/21

Training on DRD Generation & Mock Data Validation

This week, I received extensive training on the generation of Data Requirements Documents (DRD) and the processes involved in mock data validation. The key takeaways include:

Understanding DRD Generation: I gained exposure to the tools and methodologies used to generate DRDs, specifically utilizing the IA tool. This included learning about the various stages involved in creating a comprehensive and effective DRD.

Purpose and Importance: I delved into the reasons behind generating DRDs, understanding their critical role in ensuring data integrity, consistency, and compliance with regulatory standards.

Database Management and SQL Skills Enhancement

I dedicated a significant portion of this week to enhancing my database management and SQL skills, focusing on both MS SQL Server and PostgreSQL.

MS SQL Server:

Installation and Setup: Successfully installed the SQL Server Management Tool, configuring it to work seamlessly with my development environment.

Server Connection and Exploration: Navigated through the process of connecting to various databases within the server. This included running multiple SQL queries to familiarize myself with the tool’s functionalities and capabilities.

PostgreSQL:

Installation: Successfully installed PostgreSQL, setting up a robust environment for database management and development.

Server Connection Challenges: Attempted to connect to the StateFarm server, encountering challenges due to the requirement of a localhost environment. This experience highlighted the importance of environment-specific configurations.

Issue Resolution: Learned the process of raising tickets in Wiser for the StateFarm client, which is crucial for effective issue resolution and support.

Testing and Troubleshooting SQL Server Queries in PostgreSQL

An integral part of my week involved testing and troubleshooting SQL server queries in the PostgreSQL environment. Key activities and learnings include:

Collaboration and Peer Support: With the assistance of a peer, I was able to execute and test SQL queries on the PostgreSQL platform. This collaboration was instrumental in navigating initial challenges and understanding the environment better.

Error Identification and Resolution: Encountered and diagnosed errors in the output. Through detailed analysis, I identified that the root cause was an incorrect database schema specified in the query. This exercise underscored the importance of precise schema definitions and the impact of errors on query performance.

Learning about Types of Solutions Offered by AML, Fraud & Authentication Management Portfolio, FMC, Essential Cloud Offering, and Actimize Watch

AML Solutions: I explored various solutions offered for Anti-Money Laundering (AML). These solutions are crucial for detecting and preventing suspicious activities within financial transactions.

Fraud & Authentication Management: I gained insights into the portfolio of solutions designed to combat fraud and ensure robust authentication mechanisms. These solutions help in securing financial transactions and protecting customer data.

FMC (Financial Management Cloud): I learned about the essential offerings of FMC, which provide a comprehensive suite of financial management applications delivered on the cloud. These solutions help streamline financial operations and enhance efficiency.

Actimize Watch: I studied Actimize Watch, a service that provides continuous monitoring and updates on global watch lists. This helps financial institutions stay compliant with regulatory requirements and mitigate risks associated with financial crimes.

Difference between SQL Server and PostgreSQL

SQL Server: A relational database management system (RDBMS) developed by Microsoft. It is widely used in enterprise environments due to its robust performance, security features, and integration with other Microsoft products.

PostgreSQL: An open-source RDBMS known for its advanced features, compliance with SQL standards, and support for complex queries and data types. It is highly extensible and suitable for a wide range of applications.

May 23, 2024

Understanding Phases of a Project - SOP and Waterfall Methodologies

Design & Development: Focused on creating functional specifications and detailed design documents that outline the system architecture, components, and interactions.

DRD (Data Requirements Document): Developed comprehensive DRDs to define data needs, ensuring all necessary data elements and their relationships are clearly specified.

Data Validation: Implemented processes to ensure data accuracy, consistency, and integrity. This step is crucial for maintaining high-quality data throughout the project lifecycle.

ETL (Extract, Transform, Load): Executed ETL processes to integrate data from various sources, transform it to fit operational needs, and load it into the data repository.

Data Gap Analysis: Identified gaps in the existing data and devised strategies to address these gaps, ensuring complete and accurate data coverage.

System Integration Testing: Conducted thorough testing to ensure all system components interact seamlessly and perform as expected in an integrated environment.

UAT (User Acceptance Testing): Facilitated UAT to gather feedback from end-users, ensuring the system meets their requirements and is ready for deployment.

Production: Successfully transitioned the project to the production environment, ensuring it is fully operational and ready for end-users.

How to Generate DRD from IA Tool

Installation and Setup: Learned how to install the IA tool using MS Access, setting up the necessary environment for DRD generation.

Hands-on Experience: Gained practical experience in generating DRDs for SAM (Suspicious Activity Monitoring) solutions. This involved defining data elements, establishing relationships, and ensuring data quality and compliance.

May 24, 2024

Testing Updated Oracle SQL Queries in PostgreSQL

Peer Collaboration: Worked closely with a peer to run and test updated SQL queries in the PostgreSQL environment. This collaboration facilitated knowledge sharing and skill enhancement.

Error Identification and Troubleshooting: Encountered errors in the output and, with guidance from my peer, identified that the root cause was an incorrect database schema specified in the query. This experience underscored the importance of precise schema definitions and highlighted the differences between Oracle and PostgreSQL environments.