

Understanding Enterprise Data Pipeline & Architecture Report- Week 4

Name: Joori Ahmed Shareef

Company: Jedco

Department: Business Applications – Data Center 2

Duration: 30 hours

1. Objectives

This week's objective was to explore how enterprise data pipelines are designed and managed at JEDCO, particularly focusing on the source systems, integration platforms, and data quality mechanisms. Additionally, we were introduced to the foundational concepts of Dataiku, the advanced analytics platform adopted by the organization.

2. Activities Completed

- Attended internal briefing sessions explaining the end-to-end architecture of JEDCO's data warehouse system.
- Studied a detailed architecture document outlining the flow of data from source systems (SAP, AMS, Manual files) through Informatica and PostgreSQL, toward consumption platforms like Power BI and Dataiku.
- Reviewed documentation about the 25 data quality rules, categorized under accuracy, completeness, and consistency.
- Understood how SharePoint is planned to replace manual data entry to enhance automation and reliability.
- Explored the semantic layer's role in connecting the data warehouse to dashboards for business reporting.
- Prepared for the upcoming Dataiku onboarding session with the vendor team to gain hands-on exposure.

3. Key Learnings

- Understood the modular structure of JEDCO's data pipeline, involving multiple components like Informatica for ETL, PostgreSQL for staging/storage, and Power BI for visualization.
- Learned the importance of maintaining high data quality using defined rules to ensure business trust in analytics.
- Gained knowledge of the shift from manual files to automated systems like SharePoint and centralized APIs.
- Became familiar with how Dataiku will serve as the main platform for future analytics tasks and model development.

4. Supervisor Feedback

Our supervisor highlighted the strategic value of automating data flows and ensuring high-quality standards. They also emphasized the need for interns to understand not just tools but the data governance and reliability aspects behind enterprise analytics.

5. Reflection

This week deepened my appreciation of the technical and operational complexity behind enterprise analytics. Understanding architecture gave me insight into how raw data transforms into meaningful dashboards and predictive models. Although we haven't started using Dataiku yet, I feel well-prepared for the upcoming sessions thanks to the strong foundational materials.

Next Steps:

- Attend the Dataiku vendor meeting and note down key concepts from the demo.