

Capstone Requirement Document

Trendy Fabrics



Google Cloud

Table Of Contents

1. Executive Summary

1.1 Overview

1.2 Situation Appraisal

2. Business Objectives

2.1 Easy Integration

2.2 High Level Security

2.3 Scalability

2.4 Availability

2.5 Cost Efficiency

3. Requirements

3.1 Real Time Ingestion

3.2 ETL and Staging

3.3 Encryption and IAM Security

3.4 Email Automation

3.5 Application with CI/CD Enabled

3.6 Data Processing in Warehouse

3.7 Real Time Data Analytics

4. Timeline and Milestones



Google Cloud

Executive Summary

Company Overview

Trendy Fabrics is a leading textile manufacturer specializing in the production of premium textiles. With a substantial fleet of over 1000 industrial machines and a dedicated workforce exceeding 2000 employees, the company has positioned itself as a pivotal entity within the textile industry.

Current Situation Assessment

Trendy Fabrics is currently advancing its production capabilities through the successful integration of IoT solutions for real-time data acquisition from its industrial machinery. Looking ahead, the company is strategically focused on establishing a robust infrastructure optimized for sophisticated data processing, including comprehensive data analytics and the development of advanced machine learning models.



Google Cloud

Key Objectives:

Seamless Integration with Existing Systems: Facilitate the integration of new technologies, such as IoT solutions, with existing production systems to ensure a smooth and cohesive workflow.

High-Level Security: Implement robust security measures to protect the integrity and confidentiality of data collected from industrial machines, ensuring compliance with industry standards.

Scalability: Design the infrastructure to support future growth, enabling the easy expansion of IoT solutions, data analytics, and machine learning capabilities without compromising performance.

Continuous Availability: Ensure the continuous availability of critical production data and analytics tools to minimize downtime and maximize the reliability of the modernized production environment.

Cost Efficiency: Optimize the overall cost of the modernization initiative by adopting efficient technologies, reducing resource wastage, and maximizing the return on investment in data analytics and machine learning.



Google Cloud

Requirements:

Real-time Ingestion:

Develop a platform to generate data every 40 seconds from 10 machines.

Host a document database to onboard real-time data from these machines.

Extract, Transform, and Load (ETL):

Select an ETL tool to transform the data and push it into a data warehouse.

Implement a staging area to onboard raw data for future processing.

Encryption and Security:

Implement proper IAM (Identity and Access Management) rules for security.

Encrypt highly sensitive columns using a key management system.

Data Processing:

Transform the data using queries in the data warehouse to enhance performance.

Build views and stored procedures as required for the analytics platform and applications.



Requirements:

Logging and Monitoring:

Implement a comprehensive logging and monitoring system.

Report Generation and Email Automation:

Host a platform to generate reports.

Automate email delivery to send processed reports at regular intervals.

Store processed reports for future reference.

Real-time Analytics:

Collect metrics for real-time dashboarding.

Build reports using an analytics platform.

Application Development:

Develop an application to list and download processed reports from the storage bucket.

Build the application with CI/CD enabled for developer convenience.



Google Cloud

Timeline and Milestones

- Requirements Gathering and Analysis (**2 Days**)
- Infrastructure Design and Setup (**2 Days**)
- Real Time ingestion and data hosting (**3 Days**)
- ETL and Processing (**3 Days**)
- Encryption and Security(**3 Days**)
- Data Engineering in warehouse (**4 Days**)
- Report Generation and Email Automation (**4 Days**)
- Application Development and Real Time Dashboard (**5 Days**)



Communication Plan

Discussion I Understanding the Requirements (**01.07.2024 - 02.07.2024**)

Discussion II Proposal Discussion (**03.07.2024**)

Discussion III Progress and Road Block Discussion (**19.07.2024**)

Discussion IV Final Review (**29.07.2024**)



Google Cloud