Requirements Document for Intelligent Data Integration Framework

Project Name

Intelligent Data Integration Framework

Executive Summary

The Intelligent Data Integration Framework aims to automate and enhance the data integration process by leveraging artificial intelligence. This framework is designed to manage data from various sources with distinct schemas, ensuring efficient data transformation and cleansing. The use of AI in schema matching, anomaly detection, and data correction will significantly reduce manual intervention, improve data quality, and streamline ETL processes.

Tasks and Task Dependencies

Task 1: Schema Matching Automation

- **Description**: Develop AI models capable of automatically matching schemas from disparate data sources.
- **Dependencies**: Requires initial data source analysis.
- **Assigned Employees**: Cynthia Harris, Robert Stewart
- **Estimated Duration**: 4 weeks
- **Required Skills**: AI model training, schema analysis, Python programming.

Task 2: Data Transformation and ETL Automation

- **Description**: Create automated data transformation rules and streamline ETL processes using AI-generated suggestions.
- **Dependencies**: Successful completion of schema matching models.
- **Assigned Employees**: Kathleen Stewart, Sarah Jones
- **Estimated Duration**: 6 weeks
- **Required Skills**: ETL process design, AI integration, data transformation techniques.

Task 3: Data Cleansing Module Development

- **Description**: Develop a module to detect and correct anomalies, missing values, and duplicates.
- **Dependencies**: Integration with transformation rules.
- **Assigned Employees**: Susan Moore, Linda Parker
- **Estimated Duration**: 5 weeks
- **Required Skills**: Data cleansing techniques, anomaly detection, AI integration.

Task 4: User Interface Design and Implementation

- **Description**: Design a drag-and-drop interface for workflow management and real-time analytics.

- **Dependencies**: Development of core framework functionalities.
- **Assigned Employees**: Emily Collins, Jennifer Rogers
- **Estimated Duration**: 4 weeks
- **Required Skills**: UI/UX design, JavaScript, real-time data visualization.

Task 5: Cloud Integration and Deployment

- **Description**: Ensure the framework supports integration with cloud providers and on-premises systems.
- **Dependencies**: Completion of core framework development.
- **Assigned Employees**: Timothy Johnson, Rebecca Richardson
- **Estimated Duration**: 3 weeks
- **Required Skills**: Cloud computing, deployment strategies, API integration.

Task 6: Continuous Learning and Feedback Loop Implementation

- **Description**: Develop a feedback loop to improve AI models through user corrections and preferences.
- **Dependencies**: Operational user interface and core functionalities.
- **Assigned Employees**: Entire team collaboration
- **Estimated Duration**: 2 weeks
- **Required Skills**: Machine learning, user feedback analysis, iterative model improvement.

Business Rules / Constraints

- **Data Security**: Ensure compliance with data protection regulations, including GDPR and CCPA, by implementing robust security measures and encryption protocols.
- **Cloud Agnosticism**: Framework must remain cloud-agnostic, supporting seamless integration with AWS, Azure, Google Cloud, and on-premises systems.
- **Scalability**: The framework must be scalable to accommodate varying volumes of data and increasing integration demands.
- **User Accessibility**: Interface must be user-friendly, catering to users with varying levels of technical expertise, emphasizing ease of use and accessibility.

Conclusion

This document outlines a comprehensive plan to develop the Intelligent Data Integration Framework. Each task is strategically arranged to ensure a smooth workflow, with dependencies clearly identified to facilitate efficient project management. With the team's expertise and the outlined tasks, we are poised to achieve the project's objectives, delivering significant operational and technical improvements while adhering to stringent business rules and constraints.