**➢ Why is this project initiated?**

The DataHub project is important because it helps the organization to manage and use data more effectively:

* **Centralized data storage** DataHub brings together data from multiple sources into a single location, which can help ensure data consistency and reliability.
* **Improved data discovery** DataHub helps users explore and understand their data and track data lineage.
* **Better data governance** DataHub helps organizations establish data contracts and ensure compliance with regulatory requirements.
* **Improved decision-making** DataHub makes it easier to access relevant data at the right time, which can help organizations make better decisions.
* **Real-time data processing** DataHub enables organizations to analyze and act on data in near real-time.
* **Scalability** DataHub's modular design allows it to scale with data growth

 **➢ What are the current problems?**

* **Data quality**: Poor data quality, such as missing values, duplicate records, or outdated information, can lead to incorrect decisions, disrupted business processes, and undermine the new system's integrity.
* **Data loss**: Data can be accidentally deleted or corrupted during migration, which can lead to compliance failures and disrupt operations. Recovering lost or corrupted data can be costly or even impossible.
* **Security and compliance**: Data migration can pose significant security and compliance risks, especially when dealing with sensitive or regulated data. Breaches or compliance lapses can result in severe legal and financial penalties.
* **Lack of testing**: Not testing the migration process can lead to many errors and unforeseen problems. This can be due to a lack of time, resources, or poor planning.
* **Compatibility**: Software or hardware compatibility issues can arise in the new environment.
* **Operational disruptions**: Downtime and service interruptions can occur during migration.
* **User adoption**: User adoption challenges can arise.
* **Integration difficulties**: Integration difficulties can arise.
* **Unorganized data**: Data may be available in an unorganized state.
* **Inadequate data protection**: Data protection may be inadequate.
* **Incorrectly formatted data**: Data may be incorrectly formatted.
* **Lack of collaboration**: There may be a lack of collaboration among teams.
* **No process integration**: Processes may not be integrated.

**➢ With this project how many problems could be solved?**

After implementing this project, most of the issues mentioned above could be solved.

**➢ What are the resources required?**

* Software: Sources, Airflow, SqlServer, Putty, Winscp, AquaDatastudio
* Project Team : Product Owner, BA, Developers, Testers

**➢ How much organizational change is required to adopt this technology?**

The amount of organizational change required to adopt for datahub project depends on the project's scope and scale:

* Project scope and scale: A medium-scale project may take 2–4 months, while a large-scale project may take 6–12 months or longer.
* Data volume, variety, and quality: Assessing the data to be migrated helps plan resources, set timelines, and establish objectives.
* Data sensitivity: Identifying the data's format, location, and sensitivity is important.
* Migration approach: A big-bang migration involves transferring all data in a single phase.

**➢ Time frame to recover ROI?**

In this datahub project, timeframe required could be 3-4 years.

**➢ How to identify Stakeholders?**

When identifying stakeholders for a data hub project, we need to consider the following:

* Roles: Who are the key stakeholders, such as business users, data owners, IT staff, vendors, regulators, and customers?
* Responsibilities: What are the responsibilities of each stakeholder?
* Needs and concerns: What are the needs and concerns of each stakeholder?
* Impact: How will the data migration affect each stakeholder?

In the waterfall model we work closely with stakeholders to identify and document their needs. The BA's skills and responsibilities include:

* Eliciting and analyzing requirements.
* Preparing the BRD
* Clearing doubts about requirements
* Reviewing test cases
* Verifying delivery
* Reviewing change requests

Used a variety of elicitation techniques in this waterfall model project, including

* Brainstorming
* Stakeholder analysis
* Document analysis
* Interviews
* Workshops
* Observations
* Surveys and questionnaires
* Prototyping
* Interface analysis

Stakeholder analysis can help minimize risk and ensure project success

* Identify stakeholders
* Analyze stakeholders
* Prioritize stakeholders
* Plan communications
* Develop strong stakeholder relations

A RACI matrix is a tool that can help with stakeholder analysis by assigning roles and responsibilities to stakeholders. RACI stands for Responsible, Accountable, Consulted, and Informed. Here are some steps for using a RACI matrix:

* Identify tasks:
* Determine roles
* Assign RACI values
* Clarify definitions:
* Review and adjust:

As part of BA approach strategy, I usually prepare/write the following documents

* Business Requirements Document (BRD)
* Stakeholder Management Plan.
* System Requirements Specification Document (SRS)
* Functional/Process Document.

To get sign off on documents, we usually follow a process that includes:

* **Identifying stakeholders**:
* **Tracking sign-off**:
* **Conducting a sign-off meeting**:
* **Using digital signatures**:
* **Closing the loop with stakeholders**:

Client approval is an important milestone in the collaboration between a client and a contractor. Here are some tips for requesting approval from a client:

* **Be clear**: State your purpose clearly and early in your request.
* **Provide context**: Include necessary details and context.
* **Be polite**: Start with a polite greeting and address recipients professionally.
* **Set a deadline**: Specify a deadline for a response.
* **Explain the benefits**: Explain how the client will benefit from your request.
* **Use facts**: Include relevant documents and facts.
* **Conclude concisely**: End your message on a concise note.

We can use a variety of channels, including:

* **Email**:
* **Live chat**:
* **Messaging apps**:
* **Video messaging**:
* **Text messaging**:
* **Social media**:
* **Phone calls**:
* **Webinars**:
* **Forums**:
* **Chatbots**:

But, most common used channels are Email, Livechat, Webinars.

I also handled change requests by following below steps:

* **Determine the scope**
* **Analyze the impact**
* **Get approval**
* **Prepare documentation**:
* **Prioritize**
* **Communicate**
* **Get sign-off**
* **Maintain logs**

I regularly updated client on project progress by:

* **Preparing**: Before updating stakeholders, identify who they are, what they want, and how they prefer to receive information. You can use a stakeholder analysis tool to help you prioritize stakeholders and tailor your communication.
* **Planning**: Plan your presentation in advance and consider preempting questions.
* **Communicating**: Use clear, concise language, and consider using visual aids like charts, graphs, or tables. Be respectful and professional, and consider the stakeholders' culture, background, and expectations.
* **Engaging**: Ask questions, solicit feedback, and address concerns.
* **Following up**: Thank stakeholders for their time, summarize key points, and follow up on any inquiries or requests.
* **Establishing a schedule**: Set up regular meetings, such as weekly or bi-weekly, to provide updates.
* **Using a project status report:** Use a template to structure your update and include details like progress, milestones, risks, issues, and next steps.
* **Delegating**: For complex projects, delegate updates to the right operational team leads.
* **Celebrating**: Celebrate achievements and successes with stakeholders.

To take sign-off on a UAT-client project acceptance form, I can:

* **Prepare:** Gather information about the project, create a test plan, and prepare a sign-off form
* **Conduct a final UAT session**
* **Record the session:** Record the session for future reference
* **Ask for sign-off**
* **Use a sign-off form**

**Functional Specifications**

|  |  |
| --- | --- |
| Project Name | Datahub |
| Customer Name | AML |
| Project Version | V1.0 |
| Project Sponsor | FIS  |
| Project Manager | Venkateswara Rao |
| Project Initiation Date | 06/06/2025 |

**Functional Requirement specifications:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Req ID** | **Req Name** | **Req Description** | **Priority** |
| FR001 | Data Integration | Pull data from different source systems and integrate them into single system | 10 |
| FR002 | Data Migration | Migrate data from different systems to one end system called Datahub. | 10 |
| FR003 | Resources | Software requirements are databases, UNIX, IBM Datastage, JIRA. | 9 |
| FR004 | Project Team | 1 Product Owner, 1 BA, 1 Project Manager, 3 develpers, 3 Testers | 9 |
| FR005 | Sourcesystems check | Check the configuration of the source systems and set up the hub on this compatibility | 10 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Req ID** | **Req Name** | **Req Description** | **Design** | D1 | T1 | D2 | T2 | UAT |
| FR001 | Data Integration | Pull data from different source systems and integrate them into single system | Source.docx, Data should be integrated from all sources. | Done | Done | Done | Done | Done |
| FR002 | Data Migration | Migrate data from different systems to one end system called Datahub. | Migration of data should be done into Datahub without data redundancy. Data quality should be maintained, data confidentiality should be taken care. | Done | Done | Done | Done | Done |
| FR003 | Resources | Software requirements are databases, UNIX, IBM Datastage, JIRA. | all the needed softwares and applications should be installed in local machines for project team. | Done | Done | Done | Done | Done |
| FR004 | Project Team | 1 Product Owner, 1 BA, 1 Project Manager,  | done | Done | Done | Done | Done | Done |

B

 Datahub

 P66279

 V1.0

 Prathima Authur

Contents

1. Document Revisions......................................................................................................................... 3

2. Approvals...........................................................................................................................................3

3. RASCI Chart for This Document.........................................................................................................3 RASCI Chart...........................................................................................................................................3

4. Introduction......................................................................................................................................4 4.1. Business Goals..................................................................................................................... .........4 4.2. Business Objectives..................................................................................................................... 4 4.3. Business Rules....................................................................................................................... .... ..4 4.4. Background............................................................................................................................ ......4 4.5. Project Objective................................................................................................................... .......4 4.6. Project Scope................................................................................................................................5 4.6.1. In Scope Functionality................................................................................................................5 4.6.2. Out Scope Functionality.............................................................................................................5

5. Assumptions.................................................................................................................................. .5

6. Constraints.......................................................................................................................................5

7. Risks.................................................................................................................................................6 Technological Risks...............................................................................................................................6 Skills Risks.............................................................................................................................................6 Political Risks........................................................................................................................................6 Business Risks.......................................................................................................................................6 Requirements Risks..............................................................................................................................7 Other Risks......................................................................................................................................... .7

8. Business Process Overview...............................................................................................................7 8.1. Legacy System (AS-IS)....................................................................................................................8 8.2. Proposed Recommendations (TO-BE)...........................................................................................8

9. Business Requirements....................................................................................................................9

10. Appendices................................................................................................................................. 10 10.1. List of Acronyms.......................................................................................................................10 10.2. Glossary of Terms.....................................................................................................................10 10.3. Related Documents................................................................................................................. 10

1. Document Revisions

|  |  |  |
| --- | --- | --- |
| **Date** | **Version Number** | **Document Changes** |
| 06.06.2024 | 1.0 | Initial Draft |
|   |   |   |
|   |   |   |
|   |   |   |
|   |   |   |
|   |   |   |
|   |   |   |
|   |   |   |
|   |   |   |

1. Approvals

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Role**  | **Name** | **Title**  | **Signature**  | **Date** |
| Project Sponsor  | FIS |   |   |   |
| Business Owner  | Jian Chen |   |   |   |
| Project Manager | Venkateswara Rao |   |   |   |
| System Architect  | Prabhakaran B |   |   |   |
| Development Lead | Anbumozhi A |   |   |   |
| User Experience Lead  | Prasanna A |   |   |   |
| Quality Lead | Ramesh S |   |   |   |
| Content Lead | Prathima A |   |   |   |

1. RACI Chart

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Name /Position | Project Sponsor  | Business Owner  | Project Manager | System Architect  | Development Lead | User Experience Lead  | Quality Lead | Content Lead |
| FIS | C |   |   |   |   |   |   |   |
| Jian Chen |   | I |   |   |   |   |   |   |
| Venkateswara Rao |   |   | A |   |   |   |   |   |
| Prabhakaran B |   |   |   | R |   |   |   |   |
| Anbumozhi A |   |   |   |   | S |   |   |   |
| Prasanna A |   |   |   |   |   | I |   |   |
| Ramesh S |   |   |   |   |   |   | C |   |
| Prathima A |   |   |   |   |   |   |   | R |

1. Introduction
	1. Business Goals

This article is intended to help assist and standardize Data Hub organization across the AML platform. This methodology is intended to help the technical teams, navigate and create your Data Hub model and troubleshoot problems with individual data sources efficiently.

* 1. Business Objectives
* A data hub is an architecture that provides a central point for the flow of data between multiple sources and applications, enabling organizations to collect, integrate, and manage data efficiently.
* Unlike traditional data storage solutions, a data hub's purpose focuses on data integration and accessibility.
* In this project, data is brought from different sources and put into one place.
	1. Business Rules
* Prioritize business value
* Consider data sensitivity
* Use data formats that are human and machine-readable
* Work with Information Security Officers
* Ensure data use agreements are in place
* Define data governance capabilities
* Define a well-defined data management process
* Continuously monitor and improve
	1. Background

A data hub is a centralized system that collects data from various sources and makes it accessible in a unified way. It's a hybrid of a data lake and a database warehouse, and it can integrate structured, semi-structured, and unstructured data.

* 1. Project Objective
		+ Have a clear understanding of organization’s data needs and requirements
		+ Well-defined data management process which Includes data integration, data quality, data security, and data archiving.
		+ Continuous monitoring and improvement:We’ll need to continuously monitor and improve the data hub, and ensure that it remains aligned with your organization’s goals and objectives.

* 1. Project Scope

In this project, data is brought from different sources and put into one place.

* + 1. In Scope functionality
* DataHub is a modern data catalog designed to streamline metadata management, data discovery, and data governance.
* It enables users to efficiently explore and understand their data, track data lineage, profile datasets, and establish data contracts
	+ 1. Out Scope functionality
* Source system which is not compatible with datahub, this source will not be integrated.
1. Assumption
* The data will fit seamlessly into the new system
* The data transformation will be straightforward
* The data migration is a one-time event
* The current software state is documented accurately and completely.
1. Constraints
* **Data quality** issues can lead to errors, delays, and data loss. These issues can arise from incomplete, inaccurate, or duplicate data, or data that needs to be formatted.
* **Data integrity-** Loss of data integrity can occur if data validation and cleansing aren't performed before the migration.
* **Data security -** Migrating data can expose it to third parties, and the target system may be more vulnerable than the legacy system.
* **Data loss** - Data loss is a common problem that can't be avoided, even with a well-planned migration project.
* **Downtime** - Migrating data can cause system downtime, which can disrupt business operations and productivity.
* **Compatibility -** There may be compatibility issues with the new systems.
* **Integration -** There may be challenges integrating the new data with existing systems and applications.
1. Risks
* Data quality issues: These can include gaps, inconsistencies, errors, duplication, or incompleteness.
* Security vulnerabilities
* Understanding the source data
* Data mapping
* Data backup
* Application dependency mapping

Technological Risks

* + Regulatory compliance
	+ Compatibility issues:
	+ Performance degradation
	+ Hardware or software failures

Skills Risks

* + Poor communication Misalignment or unclear communication can lead to errors, delays, and misunderstandings.

To avoid this, you can:

* + Set up regular meetings and centralized communication platforms
	+ Define roles and responsibilities
	+ Align objectives
	+ Use collaboration tools

Political Risks

Political risks aren't usually a concern for data migration project

Business Risks

Data migration can pose several risks to a business, including:

* **Data loss**

Data may be accidentally deleted or corrupted during migration.

* **Data quality issues**

Common issues include missing values, duplicate records, and outdated information.

* **Data security breaches**

Inadequate migration safeguards can expose data to theft and other external threats.

* **Compliance violations**

Non-compliance with data protection laws can lead to fines and legal issues.

* **Incompatibility**

When migrating large amounts of data, it's more likely to encounter incompatible data.

* **Inadequate testing**

Insufficient testing can lead to post-migration issues that impact the new environment.

* **Data corruption**

Unwanted data migration can occur when rules and validations are set on the target system.

If not managed properly, these risks can lead to:

project delays, cost overruns, operational disruptions, customer dissatisfaction, reputational damage, and financial losses.

A thorough risk assessment can help identify potential threats and vulnerabilities

Requirements Risks

Requirements risks can be

* Misunderstanding from the Stakeholder
* Lack of gathering complete information.

Other Risks

Other risks include:

* Integration difficulties
* Regulatory compliance issues
* User adoption challenges
* Performance degradation
* Operational disruptions
* Lack of attention to detail
* Constant changes
* Poor communication
1. Business Process Overview

Data migration is the process of moving data from one storage system to another. It's a common IT activity that's often triggered by a company reorganization, merger, or acquisition. Here are some steps you can take when migrating data:

* **Pre-process**: Identify the data's location, format, and sensitivity. Plan the project's scope and size, create backups, and set up a migration tool.
* **Migration**: Execute the data migration plan.
* **Post-process**: Validate the target system, set up maintenance for the migrated system, and decommission the source system.

Here are some other things to consider when migrating data:

* **Application migration**

When an organization moves an application to a new environment, it often involves migrating the associated data. For example, a company might move from one CRM to another or implement a new HR system.

* **Data transformation**

Data often needs to be transformed as it moves from one data model to another.

* **ETL**

ETL stands for Extract, Transform, Load. It involves extracting data from sources, cleaning it, and transforming it into a format that can be understood across the enterprise.

* **Refining the scope**

This step involves filtering out excess data and defining the smallest amount of data required to run the migrated system.

**Data Hub Project Flowchart**



**Legacy system** data migration is the process of moving data, applications, and processes from outdated systems to modern platforms. This can involve updating hardware infrastructure or moving systems to a cloud infrastructure

**Proposed Recommendation**

* Business Continuity
* Improved Efficiency
* Compliance and Security
* Assess and clean your source data.

**Common Challenges in Data Migration**

* **Data Loss Risks:** Address the potential risks of data loss during migration and how to mitigate them.
* **Downtime and Disruption:** Discuss the impact of downtime during migration and strategies to minimize disruptions to business operations.
* **Data Integrity and Accuracy:** Explain the importance of maintaining data integrity and accuracy during the migration process and the consequences of failing to do so



1. Business Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Req ID** | **Req Name** | **Req Description** | **Priority** |
| FR001 | Data Integration | Pull data from different source systems and integrate them into single system | 10 |
| FR002 | Data Migration | Migrate data from different systems to one end system called Datahub. | 10 |
| FR003 | Data mapping | Use data mapping tools to outline field-level transformations and ensure alignment with the target schema | 8 |
| FR004 | Backups | Have several forms of backups in place, including a local backup and an offsite cloud backup | 7 |
| FR005 | Scalability | Performance testing evaluates the speed and efficiency of the data migration process | 6 |
| FR006 | Resources | Software requirements are databases, UNIX, IBM Datastage, JIRA. | 9 |
| FR007 | Project Team | 1 Product Owner, 1 BA, 1 Project Manager, 3 develpers, 3 Testers | 9 |

10. Appendices

 10.1. List of Acronyms

 AML – AntiMoney Laundering

BA – Business Analyst

 BRD – Business Requirement document

 ETL – Extract, Transform, Load

 FR – Functional Requirement

 RASCI – Responsible, Accountable, Supports, Consulted, Informed

 10.2. Glossary of Terms

 Data migration

Database migration

Legacy system

Data Integration

Data Mapping

10.3. Related Document

RTM

FRD