## 

## **Business & Technical Requirements for Acquire Stage**

## **1.0 Overview\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

These are sample requirements to use as a starting point in the development of your company’s business and technical requirements.

[Insert project overview description here. The text below is an example.]

The process of Leveraging Data Smartly includes identifying data science tools, technologies, methodologies, and approaches to enhance efficiency, streamline our organization’s work, and accelerate how we can discover new insights leading to actionable improvements.

Building on this process involves developing use cases (short proof of concept projects aimed at solving some of our real, data-driven challenges), each of which covers one or more of the Data Analytics Lifecycle stages of:

Data Science Lifecycle



To help you evaluate ETL (Extract-Transform-Load) technology solutions, you can leverage our recommended list of business and technical requirements.

## **2.0 Business Requirements \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Business requirements describe the activities the technology solution needs to provide (the ‘what’, not the ‘how’). These requirements provide a starting point to share and discuss with stakeholders. If there is stakeholder that owns one of the use cases that would benefit from this technology, then the owner should review these requirements and approve them. The purpose of business requirements is to document what the solution needs to be able to do and ensure there is alignment within the team.

## **3.0 Business Requirements for visualization software\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

|  |  |
| --- | --- |
| Project ID | ### |
| Name | Acquire Stage – [Project Name] |
| *Use Case* Owner | [Owner] |
| Why is a solution needed? | [Describe how data is currently collected and reviewed, and the challenges. Include potential opportunities.] |
| Why is it needed now? | [Explain why there is a current opportunity to solve this challenge.] |
| Specify needs |  |

List of Business Requirements

|  |  |
| --- | --- |
| Requirement | Description |
| 1. Data Inventory Methodology – Digital | Data in digital form needs to be inventoried and indexed.  The company has data stored in multiple locations and on multiple platforms: network drives, external hard drives, CD ROMs, hard drives, memory sticks, and email archives.  The index will help create a repository of current location details. |
| 2. Data Inventory Methodology – Hard Copy | Data that is only available in printed formats (e.g., reports, photographs) needs to be converted to a digital format.  After conversion, these datasets need to be inventoried using process from Requirement #1. |
| 3.Data Schema (e.g., method for organizing data) | A data schema is needed to design the hierarchy for how data will be organized as it is extracted.  Index and organize datasets into tables.  Provide a solution for structured and unstructured data. |
| 4.Metadata | Solution must be able to capture metadata attributes.  Ability to add tags, store characterization information, and create linkages between the datasets. |
| 5.Storage Option | Provide database solution for storing a mix of different datasets including photo libraries, multiple databases, emails, pdf file formats, and more.  The expected volume for storage is expected to be high with over 10+ years of data. |
| 6.Database Architecture | Based on the volume of data that will indexed, vendor will recommend a database structure and the attributes the database needs to support. |
| 7.Data Security | Some of the data that is being indexed is highly confidential.  There are restrictions on viewing and accessing the data that need to be controlled by an administration dashboard in the technology solution.  Security levels include:  7.1 Read and write access to raw data.  7.2 Access to modified data and reports.  7.3 Access to view datasets.  7.4 Access to view the index. |

## **4.0 Technical Requirements \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

These technical requirements will help to guide the selection and evaluation of available solutions. They describe the technical aspects such as functionality, architecture, and how the system needs to perform.

List of Technical Requirements

|  |  |
| --- | --- |
| Requirement | Description |
| 1.ETL Streaming Data Feed | The solution should be able to ingest data from streaming sources (e.g., APIs). |
| 2. ETL Digital Data | Recognize and ingest a variety of digital data formats, including (but not limited to) excel and database sources (SQL, etc.), PDF, image files, diagnostic test results, and others. |
| 3. Data Conversion | Scan or Convert non-digital data formats into digital formats. This requirements is necessary to inventory, index and make the datasets searchable. |
| 4. Data Schema | Include a schema building tool and schema templates.  Facilitate organizing datasets and creating a structure that makes sense for search and retrieval of information. |
| 5.Data Organization | Facilitates creating a Master Data List. List must be digital and easy to use and updates automatically as new datasets are extracted and stored. |
| 6.Metadata | Capture metadata attributes automatically, with the editing capabilities. Able to add tags and store characterization information. |
| 7.Data Relationships | Create dataset linkages between assets across data formats. Data is usually associated with other information and those linkages need to be stored and displayed in search results. |
| 8.Storage | Provide scalable storage with the ability to add virtual machines or private cloud storage. |
| 9.Staging Environment | Provide a location within the software where datasets are scanned for virus and malware before moving into storage.  Scan needs to occur automatically without the ability to skip this step. |

|  |  |
| --- | --- |
| Requirement | Description |
| 11.Security | All users must be authenticated with single sign-on.  Provide role based permissions to view data, edit data, ETL new data and download stored data. |
| 12.Audit Trail | Track and record user activity including datasets viewed, uploaded, downloaded, modified, and deleted.  For version control, provide a check-in and check-out function. |
| 13.Disaster Recovery | Solution should provide version control – a check-in/check-out function. |
| 14.Operating Systems | Must be compatible with Operating System Windows 10. |