



# SOFTWARE REQUIREMENT SPECIFICATIONS

[Product Name]

Version 1.0

[Company Name, Department]

## **Table of Contents**

### **1. Introduction**

- 1.1 Purpose
- 1.2 Scope
- 1.3 Definitions, Acronyms, and Abbreviations
- 1.4 References
- 1.5 Overview

### **2. Overall Description**

- 2.1 Product Perspective
- 2.2 Product Functions
- 2.3 User Classes and Characteristics
- 2.4 Operating Environment
- 2.5 Design and Implementation Constraints
- 2.6 User Documentation
- 2.7 Assumptions and Dependencies

### **3. Specific Requirements**

- 3.1 Functional Requirements
  - 3.1.1 Data Ingestion Pipelines
  - 3.1.2 Multilingual OCR
  - 3.1.3 Big Data Repository
  - 3.1.4 Search Tool
  - 3.1.5 Workflow Automation
  - 3.1.6 Text Analytics
  - 3.1.7 Geospatial Analysis Module
  - 3.1.8 Temporal Analysis Module
  - 3.1.9 Graph Analysis Module
  - 3.1.10 Predictive Module
  - 3.1.11 IMINT Module
  - 3.1.12 Widgets
  - 3.1.13 Report Generation
  - 3.1.14 AI/LLM Model
- 3.2 Non-Functional Requirements
  - 3.2.1 Performance Requirements
  - 3.2.2 Security Requirements
  - 3.2.3 Usability Requirements
  - 3.2.4 Reliability Requirements
  - 3.2.5 Maintainability Requirements
  - 3.2.6 Scalability Requirements
- 3.3 Interface Requirements
  - 3.3.1 User Interfaces
  - 3.3.2 Hardware Interfaces
  - 3.3.3 Software Interfaces
  - 3.3.4 Communication Interfaces

### **4. Quality Attributes**

### **5. Other Non-Functional Requirements**

### **6. Other Requirements**

### **7. Appendix**

# 1. Introduction

## 1.1 Purpose

This document outlines the software requirements for the "Intelligence Fusion Center" (IFC). The primary purpose of the IFC is to establish a centralized platform capable of ingesting, processing, and analyzing diverse data from multiple sources, including both structured and unstructured formats. The goal is to transform this raw data into actionable intelligence, supporting informed decision-making and enhancing operational efficiency for various stakeholders. This system aims to streamline data management, improve analytical capabilities, and provide comprehensive insights through advanced visualization and reporting tools.

## 1.2 Scope

This SRS covers all functional and non-functional requirements for the Intelligence Fusion Center, including data ingestion, processing, search, case management, visualization, reporting, and predictive analytics.

The Intelligence Fusion Center will serve as a comprehensive solution for managing and analyzing large volumes of disparate data.

**Data Sources:** The system will be designed to handle a wide array of data sources, including:

- Excel files
- Microsoft Access Databases
- SQL Databases
- Scanned and digital copies of documents (structured, semi-structured, and unstructured data)

**Key Functionalities:** The IFC will encompass the following core functionalities:

- **Data Ingestion and Collection:** Automated and manual mechanisms for bringing data into the system.
- **Data Processing:** Capabilities for cleaning, transforming, and enriching raw data, including multilingual Optical Character Recognition (OCR).
- **Unified Data Warehousing:** A central repository that compiles and correlates diverse data formats into a unified, queryable database with a comprehensive query catalog.
- **Data Visualization:** Presentation of insights through various formats such as tables, timelines, graphs, interactive dashboards, and static reports.
- **Advanced Analytics:** Implementation of predictive analytics, automated reporting, and anomaly detection to uncover patterns and identify unusual activities.
- **Intelligence Generation:** Leveraging AI/LLM models for enhanced search and summarization of key information.

## 1.3 Definitions, Acronyms, and Abbreviations

Acronym/Term	Definition
AI	Artificial Intelligence

ACL	Access Control List
DFD	Data Flow Diagram
GDPR	General Data Protection Regulation
GIS	Geographic Information System
HTML	HyperText Markup Language
IFC	Intelligence Fusion Center
ISO/IEC	International Organization for Standardization / International Electrotechnical Commission
LLM	Large Language Model
NLP	Natural Language Processing
OCR	Optical Character Recognition
PDF	Portable Document Format
RBAC	Role-Based Access Control
REST API	Representational State Transfer Application Programming Interface
RPA	Robotic Process Automation
SRS	Software Requirements Specification
SQL	Structured Query Language
UI	User Interface
XML	Extensible Markup Language

## 1.4 References

- IEEE Std 830-1998, IEEE Recommended Practice for Software Requirements Specifications.
- ISO/IEC 12207:2017, Systems and software engineering — Software life cycle processes.

## 1.5 Overview

The remainder of this SRS document details the requirements for the Intelligence Fusion

Center. Section 2 provides a general description of the product, its functions, user characteristics, and overall constraints. Section 3 elaborates on the specific functional and non-functional requirements, including system architecture, data models, and external interfaces. Finally, Section 4 includes supporting information such as use cases, data flow diagrams, and UI wireframe descriptions, linking them to the defined requirements.

## 2. Overall Description

### 2.1 Product Perspective

The Intelligence Fusion Center is a standalone system designed to integrate with existing organizational data sources. It is not dependent on other major systems for its core functionality but will provide interfaces for potential future integration with third-party tools or other enterprise systems. The IFC will act as a central hub for intelligence operations, consolidating data that is currently dispersed across various formats and locations.

### 2.2 Product Functions

The primary functions of the IFC include:

- **Automated Data Ingestion:** Securely importing data from diverse sources.
- **Intelligent Data Processing:** Cleaning, transforming, and enriching data, including advanced multilingual OCR and text analytics.
- **Unified Data Storage:** Creating a cohesive, searchable, and queryable repository for all ingested data.
- **Advanced Search Capabilities:** Enabling users to find specific information quickly and effectively through various search methodologies.
- **Workflow Management:** Automating and managing the flow of information processing and report generation.
- **Comprehensive Data Visualization:** Presenting complex data in intuitive and interactive visual formats (GIS maps, timelines, graphs, dashboards).
- **Predictive Detection:** Identifying trends, forecasting future events, and unusual patterns in the data.
- **Automated Report Generation:** Producing customizable reports in multiple formats.
- **AI-Powered Intelligence:** Utilizing LLMs for advanced search, summarization, and analysis of large text datasets.
- **User Management and Security:** Authentication, Authorization, and auditing.

### 2.3 User Classes and Characteristics

The primary users of the Intelligence Fusion Center will include:

- **Data Analysts:** Users with a strong understanding of data analysis techniques, requiring access to raw and processed data, advanced search, and analytical tools. They will utilize the system for in-depth investigations and intelligence generation.
- **Intelligence Officers:** Users focused on actionable insights, requiring dashboards, reports, and summarized intelligence. They will primarily interact with the visualization and reporting modules.
- **System Administrators:** Users responsible for system configuration, user management, data source integration, and monitoring system performance. They will require administrative interfaces.

All users are expected to have basic computer literacy. Training will be provided for specific functionalities.

## 2.4 Operating Environment

- Web browser (Chrome, Firefox, Edge)
- Server operating system (Ubuntu Linux, Windows)
- Database server (Elasticsearch, Neo4J, MySQL databases)
- Big Data processing framework
- Selfhosted VM environment Optional.

## 2.5 Design and Implementation Constraints

- Adherence to security standards.
- The system should implement robust Role-Based Access Control (RBAC).
- Use of scalable and fault-tolerant architecture.
- Use of open source where possible.
- Real-time queries should result in nominal time period.
- The system must effectively handle structured, semi-structured, and unstructured data from various sources (Excel, Access DBs, SQL DBs, scanned/digital documents).
- OCR and text analytics should support multiple languages including English, Hindi, and Mandarin.

## 2.6 User Documentation

- User manual
- API documentation
- Training materials

## 2.7 Assumptions and Dependencies

- **Data Quality:** It is assumed that while raw data may be messy, it will be in a format that allows for automated parsing and processing after initial ingestion. Significant data quality issues in source systems may impact processing efficiency.
- **Infrastructure Availability:** The necessary hardware and network infrastructure will be available and appropriately scaled to support the system's performance and scalability requirements.
- **User Training:** Users will receive adequate training to effectively utilize the system's features.
- **API Access:** For third-party integrations, it is assumed that necessary API keys and access permissions will be provided by external systems.

# 3. Specific Requirements

## 3.1 Functional Requirements

This section details the specific functionalities the Intelligence Fusion Center must provide.

### 3.1.1 Data Ingestion Pipelines (FR-ING-001)

- **FR-ING-001.1:** The system shall provide dedicated, configurable pipelines for ingesting data from Excel files.
- **FR-ING-001.2:** The system shall provide dedicated, configurable pipelines for ingesting data from Microsoft Access Databases.
- **FR-ING-001.3:** The system shall provide dedicated, configurable pipelines for ingesting data from SQL Databases.
- **FR-ING-001.4:** The system shall provide dedicated, configurable pipelines for ingesting scanned documents.
- **FR-ING-001.5:** The system shall provide dedicated, configurable pipelines for ingesting digital document files (e.g., PDF, Word).
- **FR-ING-001.6:** The system shall support only on-demand data ingestion for all supported sources.

### 3.1.2 Multilingual OCR (FR-OCR-001)

- **FR-OCR-001.1:** The system shall perform Optical Character Recognition (OCR) on scanned documents to extract text.
- **FR-OCR-001.2:** The OCR module shall accurately recognize and extract text in English.
- **FR-OCR-001.3:** The OCR module shall accurately recognize and extract text in Hindi.
- **FR-OCR-001.4:** The OCR module shall accurately recognize and extract text in Mandarin.
- **FR-OCR-001.5:** The system shall allow users to review and correct OCR errors.

### 3.1.3 Big Data Repository (FR-BDR-001)

- **FR-BDR-001.1:** The system shall compile and store data from all ingested sources (Excel, Access, SQL, scanned/digital documents) in a unified database.
- **FR-BDR-001.2:** The repository shall support storing structured, semi-structured, and unstructured data formats.
- **FR-BDR-001.3:** The system shall automatically co-relate different data sources based on predefined or configurable relationships.

### 3.1.4 Search Tool (FR-SRCH-001)

- **FR-SRCH-001.1:** The system shall import raw data into a search index for efficient retrieval.
- **FR-SRCH-001.2:** The system shall offer form-based search, allowing users to specify criteria in structured fields.
- **FR-SRCH-001.3:** The system shall offer advanced search capabilities supporting Boolean logic (AND, OR, NOT).
- **FR-SRCH-001.4:** The system shall support proximity search, allowing users to find terms within a specified distance of each other.
- **FR-SRCH-001.5:** The system shall support fuzzy search, enabling retrieval of results even with minor misspellings or variations in search terms.
- **FR-SRCH-001.6:** Search results shall be presented clearly, indicating the source and relevance.

### 3.1.5 Workflow Automation (FR-WFA-001)

- **FR-WFA-001.1:** The system shall support the definition and execution of automated workflows for processing ingested information.
- **FR-WFA-001.2:** Workflows shall allow for multi-level processing stages (e.g., ingestion, enrichment, analysis, reporting).
- **FR-WFA-001.3:** The system shall enable users to configure rules and triggers for workflow execution.
- **FR-WFA-001.4:** The system shall provide a mechanism for users to review and approve intermediate and final outputs of workflows.
- **FR-WFA-001.5:** The system shall generate reports automatically based on defined workflow outcomes.

### 3.1.6 Text Analytics (FR-TA-001)

- **FR-TA-001.1:** The system shall incorporate text analytics capabilities as part of the data pre-processing stage.
- **FR-TA-001.2:** The text analytics module shall perform text classification to categorize documents or text segments.
- **FR-TA-001.3:** The text analytics module shall support dictionary augmentation, allowing users to add custom terms and synonyms.
- **FR-TA-001.4:** The text analytics module shall perform entity extraction to identify and categorize key entities (e.g., units, locations, dates) from text.
- **FR-TA-001.5:** The extracted entities and classifications shall be stored alongside the data for analytical querying.

### 3.1.7 Geospatial Analysis Module (FR-GIS-001)

- **FR-GIS-001.1:** The System should have an inbuilt/ integrated GIS that shall display geospatial data (inputs, reports, analysis) on an interactive GIS dashboard.
- **FR-GIS-001.2:** The GIS dashboard shall support various map layers and overlays.
- **FR-GIS-001.3:** Users shall be able to filter and interact with geospatial data directly on the map.
- **FR-GIS-001.4:** The system shall support plotting custom points, events, areas, and routes on the map. On click it should also show associated information of that events.
- **FR-GIS-001.5:** The system has ability to export geo-referenced data in a format compatible with OGC.
- **FR-GIS-001.6:** The system should generate and visualise the heat map, cluster map and charts.
- **FR-GIS-001.7:** The system should allow creation of geo-polygon boundaries for filtered searches.
- **FR-GIS-001.8:** The system should allow creation of timeline layer on top of the geo-layer.
- **FR-GIS-001.9:** The system should allow import and export of GIS data such as shp & KML files geographics data in an excel/csv file.

### 3.1.8 Temporal Analysis Module (FR-TIM-001)

- **FR-TIM-001.1:** The system shall provide a timeline analysis module.
- **FR-TIM-001.2:** The timeline module shall allow multiple events from different data sources to be plotted on the same chronological timeline.
- **FR-TIM-001.3:** Users shall be able to zoom, pan, and filter events on the timeline.



- **FR-TIM-001.4:** The timeline shall display event details upon selection.

### **3.1.9 Graph Analysis Module (FR-GRPH-001)**

- **FR-GRPH-001.1:** The system shall represent relationships between data nodes in an interactive graph visualization.
- **FR-GRPH-001.2:** The graph module shall visualize reports and analytical findings.
- **FR-GRPH-001.3:** Users shall be able to explore connections and expand nodes.
- **FR-GRPH-001.4:** The graph visualization shall support different layout methods.

### **3.1.10 Predictive Module (FR-PAN-001)**

- **FR-PAN-001.1:** Predictive Generation Capability
  - The system shall provide the ability to generate predictions based on analyzed trends and patterns.
  - The system shall support predictive generation for user-defined "Use Cases" (specific events).
- **FR-PAN-001.2:** The system shall provide the ability to perform pattern recognition on historical data for user-defined "Use Cases."
- **FR-PAN-001.3:** The system shall allow users to define and select specific "Use Cases" for predictive generation and pattern recognition.
- **FR-PAN-001.4:** The system's prediction module shall utilize historical data from the past three years.
- **FR-PAN-001.5:** The system shall identify trends based on the analyzed historical data.
- **FR-PAN-001.6:** The system shall make predictions for specific scenarios, including but not limited to:
  - Likely hotspots for infiltration.
  - Units which may go for training or move.
  - Likely hotspots (concentration) of radars.
- **FR-PAN-001.7:** The predictions shall be based on a mathematical model built on probabilities.
- **FR-PAN-001.8:** The predictive module shall achieve a minimum accuracy of 30% at the SSCT (System/Software/Solution/System Component Testing) Stage.

### **3.1.11 IMINT Module (FR-IMT-001)**

- **FR-IMT-001.1:** The system shall enable the user to ingest relevant data from the IMINT into a form-based collation system. For trial purposes, the system shall allow the user to provide IMINT data in their existing IMINT format.
- **FR-IMT-001.2:** The system shall provide a dedicated dashboard for the user (duly standardized by user) for ingestion of IMINT from IFC workstation of IIT.

### **3.1.12 Widgets (FR-WGT-001)**

- **FR-WGT-001.1:** The Intelligence Fusion Center shall offer a variety of customizable widgets for display on dashboards.
- **FR-WGT-001.2:** Widgets shall include, but not be limited to, summary statistics, and charts (bar, pie, line).

### **3.1.13 Report Generation (FR-RPT-001)**

- **FR-RPT-001.1:** The system shall generate output reports based on user-defined criteria and workflow outcomes.
- **FR-RPT-001.2:** Reports shall be generatable in Microsoft Word (.docx) format.
- **FR-RPT-001.3:** Reports shall be generatable in Microsoft Excel (.xlsx) format.
- **FR-RPT-001.4:** Reports shall be generatable in Portable Document Format (PDF).
- **FR-RPT-001.5:** Reports shall be generatable in Extensible Markup Language (XML) format.
- **FR-RPT-001.6:** Reports shall be generatable in HyperText Markup Language (HTML) format.
- **FR-RPT-001.7:** The system shall provide a user interface for managing generated reports, including viewing, and downloading.

### 3.1.14 AI/LLM Model (FR-LLM-001)

- **FR-LLM-001.1:** The system shall incorporate a Large Language Model (LLM) to enhance search capabilities.
- **FR-LLM-001.2:** The system shall be capable of summarizing user-specified events from ingested data.
- **FR-LLM-001.3:** The system shall be capable of summarizing user-specified profiles (e.g., units, locations, equipment).
- **FR-LLM-001.4:** Users shall be able to select multiple events and/or documents for summarization by the system.
- **FR-LLM-001.5:** The system shall provide natural language search capabilities, allowing users to ask questions and receive relevant answers from the data.
- **FR-LLM-001.6:** The system should allow report generation for the user from data.
- **FR-LLM-001.7:** The system should be capable to translate any text into different languages including Hindi, Chinese, Hebrew, Arabic etc.

## 3.2 Non-Functional Requirements

This section describes the quality attributes and general constraints that the system must adhere to.

### 3.2.1 Security (NFR-SEC-001)

- **NFR-SEC-001.1:** The system shall comply with general data processing and privacy.
- **NFR-SEC-001.2:** The system shall implement Role-Based Access Control (RBAC) to restrict access to data and functionalities based on user roles and permissions.
- **NFR-SEC-001.3:** The system shall provide robust authentication mechanisms for user login.
- **NFR-SEC-001.4:** The system shall log all access attempts and data modifications for auditing purposes.

### 3.2.2 Usability (NFR-USB-001)

- **NFR-USB-001.1:** The user interface shall be intuitive and easy to navigate for all user roles.
- **NFR-USB-001.2:** The system shall provide clear feedback to users on the status of operations (e.g., ingestion progress, query execution).
- **NFR-USB-001.3:** Error messages shall be clear, concise, and provide actionable guidance.

- **NFR-USB-001.4:** The system shall offer customizable dashboards and reporting options to cater to different user preferences.

### 3.2.3 Maintainability (NFR-MNT-001)

- **NFR-MNT-001.1:** The system shall be designed with a modular architecture to facilitate easier updates, bug fixes, and feature enhancements.
- **NFR-MNT-001.2:** The system shall provide comprehensive logging and monitoring capabilities to assist in troubleshooting.

### 3.2.4 Reliability (NFR-REL-001)

- **NFR-REL-001.1:** The system is highly available.
- **NFR-REL-001.2:** The system shall implement encryption of data at rest and transit must be ensured with encryption algorithms.
- **NFR-REL-001.3:** The system shall gracefully handle errors and failures, minimizing impact on user operations.
- **NFR-REL-001.4:** Backup of user data shall be taken every week/month securely.
- **NFR-REL-001.5:** The system shall monitor all API endpoints periodically.

### 3.2.5 Scalability (NFR-SCL-001)

- **NFR-SCL-001.1:** The system architecture shall allow for horizontal & vertical scaling of microservices and Elasticsearch.

## 3.3 Interface Requirements

### 3.3.1 User Interfaces

The system will provide a comprehensive web-based user interface designed for ease of use and efficient interaction.

- **UI1:** The system shall provide a web-based user interface.
- **UI2:** The UI shall be responsive and adapt to different screen sizes.
- **UI3:** The UI shall provide a consistent look and feel across all modules.
- **UI4:** Data Ingestion
  - The application should provide an interface to import data from multiple sources such as excels, databases and documents into the Big Data Repository.
- **UI5:** Review & Approvals UI:
  - Dedicated section for reviewing generated reports and processed information.
  - Workflow-driven approval processes for intelligence outputs.
  - Ability to add comments and feedback during review.
- **UI7:** Data Visualization UI:
  - GIS Dashboard: Interactive map displaying geospatial data, with filtering and layering options.
  - Timeline Analysis: Interactive timeline for plotting and exploring multiple events chronologically.
  - Graph Analysis: Dynamic graph visualization for exploring relationships between

- data entities.
  - Dashboards: Customizable dashboards with various widgets (charts, summaries).
- **UI8: Search UI:**
  - Intuitive interface for form-based, advanced, proximity, and fuzzy searches.
  - Presentation of search results with relevance ranking and source information.
- **UI9: Report Generation UI:**
  - Interface to define report parameters and select output formats.
  - Access to generated reports for viewing and downloading.
- **UI10: LLM Interaction UI:**
  - Interface to select documents/events/profiles for summarization.
  - Input field for natural language queries to the LLM.
  - Display of LLM-generated summaries, answers, and translations.
  - Option to download the data reports for the user.

### **3.3.2 Hardware Interfaces**

- HI1: The system shall support standard server hardware configurations.
- HI2: The system shall support integration with external devices through peripheral ports.

### **3.3.3 Software Interfaces**

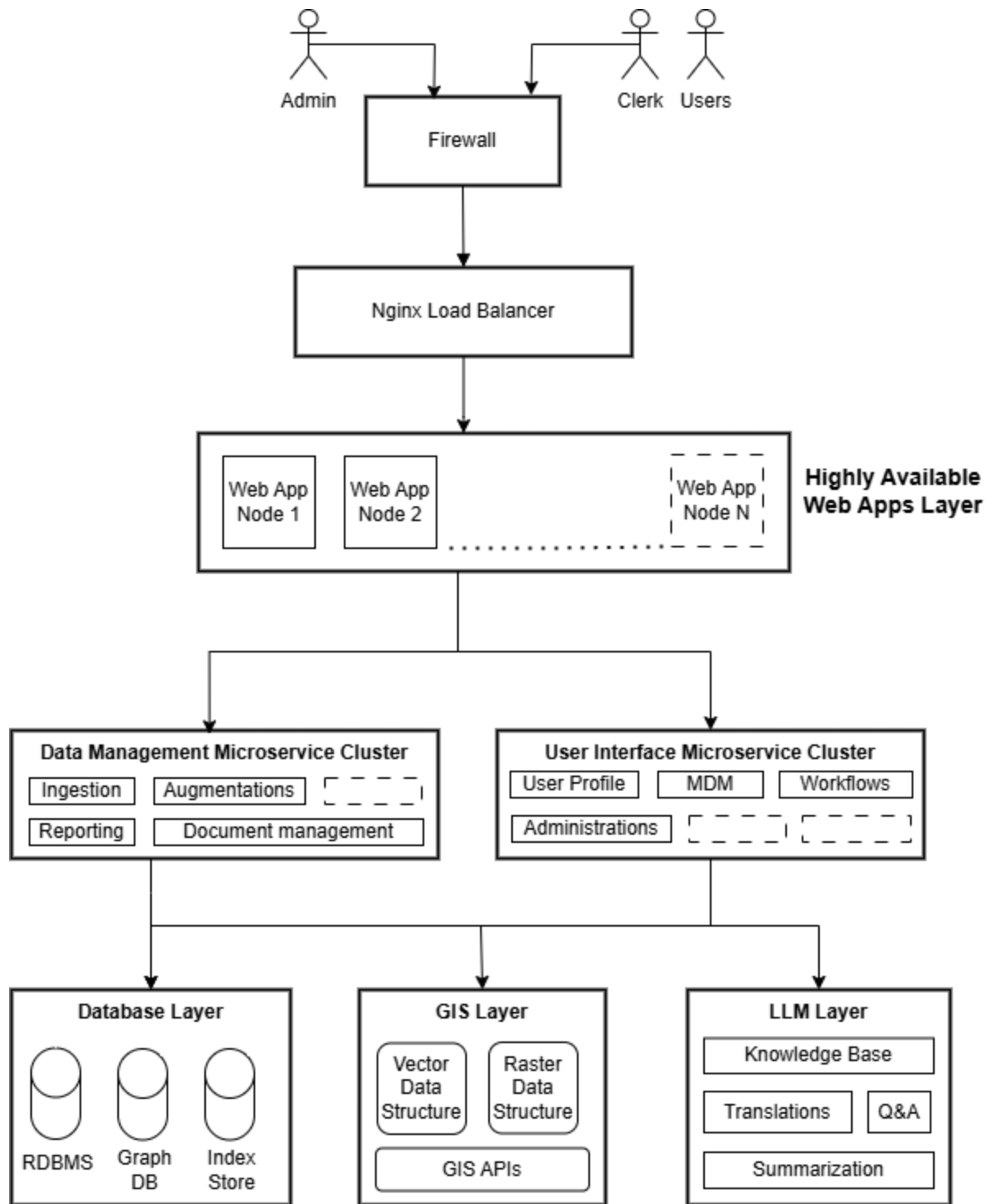
- SI1: The system shall provide APIs for integration with other systems.
- SI2: The system shall support integration with various database systems and data sources.
- SI3: Provision of secured authentication API to access data from IFC system.

### **3.3.4 Communication Interfaces**

- CI1: The system shall support secure communication protocols (HTTPS).
- CI2: The system shall support data exchange in standard formats (JSON, XML, CSV).

## **3.4 System Architecture**

The proposed system architecture for the Intelligence Fusion Center will be modular, leveraging modern technologies to ensure scalability, performance, and maintainability.



### 3.5 Data Model

The system will define clear schemas for raw data, processed data, and intelligence outputs to ensure consistency and facilitate efficient querying and analysis.

### 3.5.1 Raw Data Schema

This schema represents the data as it is ingested from source systems, retaining its original structure as much as possible.

#### Example (Conceptual):

- **Excel Data:**
  - filename: String (e.g., "unit2\_intell\_report.xlsx")
  - sheet\_name: String (e.g., "Sheet1")
  - row\_number: Integer
  - column\_headers: Array of Strings
  - row\_data: JSON Object (key-value pairs representing cell data)
  - ingestion\_timestamp: Timestamp
- **Access DB / SQL DB Data:**
  - database\_name: String
  - table\_name: String
  - record\_id: String/Integer (Primary Key from source)
  - field\_data: JSON Object (key-value pairs representing column data)
  - ingestion\_timestamp: Timestamp
- **Scanned/Digital Document Data:**
  - document\_id: String (Unique identifier)
  - document\_type: String (e.g., "PDF", "Word Document")
  - original\_filename: String
  - file\_path\_or\_url: String (Location of the original file)
  - ingestion\_timestamp: Timestamp
  - metadata: JSON Object (e.g., {'author': 'Ramesh', 'creation\_date': '2023-01-15'})

### 3.5.2 Processed Data Schema

This schema represents data after cleansing, transformation, OCR, and initial text analytics. It is optimized for storage in the unified data warehouse and for search.

#### Example (Conceptual):

- **Unified Record (Event entity):**
  - HashId: String (System-generated unique ID)
  - EventDate: String (e.g., '2024-05-10')
  - SourceCountry: String (e.g., 'China')
  - TargetCountry: String (e.g., 'Nepal')
  - EventType: String (e.g., 'Movement')
  - EMFLAG: Long (e.g., 56783568)
  - Latitude: String (e.g., '123.456')
  - Longitude: String (e.g., '321.654')
  - Unit: String (e.g., 'Unit1a')
  - Addresses: String (e.g., 'PO box 11, Pitampura, new delhi')
  - EventId: String (e.g., '123456789')

### 3.5.3 Intelligence Output Schema

This schema represents the actionable insights, summaries, and reports generated by the system, often derived from LLM processing or advanced analytics.

### Example (Conceptual):

- **Intelligence Report:**
  - report\_id: String (Unique ID)
  - report\_title: String
  - generation\_timestamp: Timestamp
  - generated\_by\_user\_id: String (User who initiated generation or workflow)
  - summary\_text: String (LLM-generated summary or analytical summary)
  - key\_findings: Array of Strings
  - associated\_record\_ids: Array of Strings (IDs of processed records used for report)
  - report\_type: String (e.g., "Event Summary")
  - status: String (e.g., "Draft", "Under Review", "Approved", "Published")
  - output\_formats\_available: Array of Strings (e.g., "PDF", "Word", "HTML")
  - geospatial\_context: JSON Object (Relevant locations for the intelligence)
  - temporal\_context: JSON Object (Relevant timeframes for the intelligence)

## 4. Quality Attributes

- **Performance:** The system shall be responsive and efficient.
- **Security:** The system shall protect sensitive data from unauthorized access.
- **Usability:** The system shall be easy to use and understand.
- **Reliability:** The system shall be available and fault tolerant.
- **Maintainability:** The system shall be easy to maintain and update.
- **Scalability:** The system shall be able to handle increasing data volumes and user loads.

## 5. Other Non-Functional Requirements

- Training requirements.
- Deployment requirements.
- Support requirements.

## 6. Other Requirements

- Data migration requirements.
- Data archival requirements.

## 7. Appendix:

- **Appendix A: Use Cases:** Detailed descriptions of how users will interact with the system.
- **Appendix B: Data Dictionary:** Definitions of all data elements used in the system.
- **Appendix C: Glossary:** Definitions of all technical terms used in the document.
- **Appendix D: Acceptance Criteria:** Specific, measurable, achievable, relevant, and time-bound (SMART) criteria for each requirement.
- **Appendix E: Risk Analysis:** Identification and mitigation strategies for potential risks.
- **Appendix F: Traceability Matrix:** A matrix that maps requirements to design elements, code modules, and test cases.
- **Appendix G: Business Rules:** Clearly stated business rules that the system must enforce.
- **Appendix H: Deployment Diagram:** A diagram showing the physical deployment of the software.
- **Appendix I: Data Flow Diagram:** A diagram showing the flow of data through the

system.

- **Appendix J: State Transition Diagram:** A diagram showing the interactions between objects in a sequence.
- **Appendix K: Sequence Diagram:** A diagram showing the interactions between objects in a sequence.
- **Appendix L: Test cases:** A list of test cases that must be executed.



# Appendix A: Use Cases

This section provides examples of use cases, illustrating how different user roles will interact with the system to achieve their goals

## Data Ingestion Pipelines (FR-ING-001)

### Use Case: Configure and Execute On-demand Data Ingestion

**Goal:** To allow a user to set up and initiate an on-demand data ingestion pipeline from various source types.

**Actors:** Clerk, System User

#### Preconditions:

- The system is operational.
- User has appropriate permissions to configure and manage data ingestion pipelines.
- Data sources (Excel, Access, SQL, scanned/digital documents) are accessible by the system.

#### Main Flow:

1. The user navigates to the "File Work" module and click on the 'Create' button for data ingestion. Add the mandatory fields like file, name, report date, file type for the datasets (file or databases) you wish to ingest.
2. **Configure Pipeline:**
  - **For File-based sources:** The user specifies the file path(s) or directory, file naming conventions.
  - **Select Source Type:** The user chooses the data source type from a list:
    - Excel File (FR-ING-001.1)
    - Microsoft Access Database (FR-ING-001.2)
    - Access Database (FR-ING-001.3)
    - Scanned Documents (FR-ING-001.4)
    - Digital Document Files (e.g., PDF, Word) (FR-ING-001.5)
  - **For Database sources (SQL):** The user provides connection details (server address, database name, credentials), and defines any filtering criteria needed.
3. **Execute Ingestion:**
  - The user clicks "submit" to initiate the ingestion process immediately.
  - The system implicitly maps the ingested data to the Big Data Repository.
4. **Ingestion Status:** The system displays a fadeout message with ingestion results.

5. **Address Errors (if any):** If failures occur, the user can review the logs, identify the root cause, and take corrective actions (e.g., correct source data) and re-run the ingestion.

**Alternative Flows:**

- **Ingestion fails due to connectivity issues:** The system logs a connection error, and the user is notified.
- **Ingestion fails due to data format issues:** The system logs data parsing errors, and the user is notified.

**Postconditions:**

- Data from the specified source is ingested into the system's Big Data Repository.
- 

### 3.1.2 Multilingual OCR (FR-OCR-001)

#### Use Case: Perform Multilingual OCR on Scanned Documents

**Goal:** To accurately extract text from scanned documents in multiple languages (such as English, Hindi, Mandarin) for further processing and searchability.

**Actors:** System (automated process, triggered by FR-ING-001.4), Data Reviewer

**Preconditions:**

- A scanned document image has been ingested into the system (via FR-ING-001.4).
- The OCR module is integrated and operational.

**Main Flow:**

1. **OCR Trigger:** Upon successful ingestion of a scanned document image, the system automatically triggers the OCR module.
2. **Language Detection (Implicit/Configurable):** The OCR module (or an upstream component) attempts to detect the language(s) present in the document. Alternatively, the ingestion pipeline configuration might allow specifying the expected language.
3. **Text Extraction:** The OCR module processes the documents pixel by pixel, recognizing characters and converting them into machine-readable text.
4. **Language-Specific Processing:** Based on detected or configured language, the OCR module applies language-specific recognition models (e.g., English, Hindi, Mandarin character sets and linguistic rules).
5. **Text Output Generation:** The extracted text is generated and associated with the original scanned document. This may include positional information for each recognized word/character.
6. **OCR Error Review (Optional):**
  - A Data Reviewer can access the document in a review interface, displaying both the original image and the extracted text.
  - The Data Reviewer manually corrects any inaccuracies in the extracted text.

7. **Store OCR Output:** The corrected (or uncorrected) OCR text is stored in the Big Data Repository, linked to the original document image.

**Alternative Flows:**

- **Incorrect language detection:** If the OCR module incorrectly detects the language, leading to poor extraction, a Data Reviewer can manually override the extracted text.

**Postconditions:**

- Scanned document image has associated extracted text.
  - Extracted text is stored in the Big Data Repository.
  - OCR errors are minimized, or a mechanism for correction is provided.
- 

### 3.1.3 Big Data Repository (FR-BDR-001)

#### Use Case: Manage and Query Unified Data Repository

**Goal:** To store, organize, and allow querying of diverse data types from multiple sources in a unified manner.

**Actors:** Clerk, Data Analyst, System

**Preconditions:**

- Data has been successfully ingested from various sources (Excel, Access, SQL, scanned/digital documents) into the system.

**Main Flow:**

1. **Data Compilation (System):** As data is ingested (FR-ING-001) and processed (e.g., OCR, FR-OCR-001), the system automatically compiles and stores it in the Big Data Repository.
2. **Data Co-relation (System/Clerk):**
  - **Automatic:** The system applies pre-defined or configured rules to identify and establish relationships between different data sources (e.g., linking event ID across Excel and SQL data).
3. **Data Type Handling (System):** The repository seamlessly handles and stores structured data (from databases, Excel), semi-structured data (e.g., JSON extracted from documents), and unstructured text (from OCR, digital documents) while maintaining their integrity and relationships.

**Alternative Flows:**

- **Query execution error:** If a query is malformed or targets non-existent data, the system provides an informative error message.

**Postconditions:**

- All ingested data is centrally stored and accessible.
- Relationships between disparate data sources are established.

---

### 3.1.4 Search Tool (FR-SRCH-001)

#### Use Case: Search and Retrieve Information

**Goal:** To enable users to efficiently search and retrieve relevant information from the unified data repository using various search capabilities.

**Actors:** System User, Data Analyst

**Preconditions:**

- Data has been ingested and stored in the Big Data Repository (FR-BDR-001).
- Data has been imported into the elastic search index.

**Main Flow:**

1. **Search Indexing (System):** Upon data ingestion or update, the system automatically imports relevant data into a search index for optimized retrieval.
2. **Access Search Interface:** The user navigates to the "Search" module.
3. **Global Search (360 search):**
  - To enable user search specific keywords or group of keywords to get all events related to those keywords.
4. **Search by Parameter (Basic Search):**
  - The user sees a search form with structured fields representing common data attributes (e.g., document, events, allorbat).
  - The user enters criteria into one or more fields.
  - The user clicks "Search."
5. **Search by Parameter (Advanced Search):**
  - The user selects the "Advanced Search" option.
  - Any of these
  - All of these
  - None of these
  - **Boolean:** The user enters search terms using Boolean operators (AND, OR, NOT).
  - **Proximity:** The user enters terms and specifies a maximum word distance.
  - **Fuzzy:** The user enters a search term with a fuzzy indicator (e.g., "receipt~" for "receipt").
  - The user clicks "Search".
6. **Display Search Results:** The system presents a list of search results as tree-map.
  - A tree map indicates the searched keywords present in all indexes with the record volume in repository

- Each result clearly indicates its source (e.g., "Excel: unit2\_intell.xls," "Document: report.docx").
- 7. **Refine/Sort Results:** The user can further refine results using filters (e.g., by date, source, document type) or sort them by relevance, date, or other attributes.
- 8. **View Full Document/Record:** The user clicks on a search result to view the full document or record from the Big Data Repository.

**Alternative Flows:**

- **No results found:** The system informs the user that no results match the criteria.
- **Permission-based results:** The system only displays results that the logged-in user is authorized to view.

**Postconditions:**

- User can efficiently find relevant information across the entire data repository.
- Search results are accurate, relevant, and clearly presented.

### 3.1.5 Workflow Automation (FR-WFA-001)

#### Use Case: Define and Execute Automated Workflows

**Goal:** To automate complex data processing tasks by defining multi-stage workflows, including review and approval mechanisms.

**Actors:** Clerk, System, Workflow Approver

**Preconditions:**

- Data has been ingested and is available in the Big Data Repository.
- Required processing modules (e.g., cleansing, enrichment, analysis) are available.

**Main Flow:**

1. **Define Workflow:**
  - The Clerk accesses the "File Work" interface.
  - The Clerk successfully ingest the multiple files into the system.
  - The Clerk will forward the uploaded source file to second higher designation.
  - Second designated user chooses relevant data from it, to generate the reports and mark it to the approver.
  - The approver may recommend edits, or forward input as is to higher designation.
2. **Workflow Execution (System/Trigger):**
  - When a defined trigger occurs (e.g., new data ingested, or approval submitted), the system automatically initiates the workflow.
  - The system executes each stage sequentially, passing the output of one stage as the input to the next.
  - During execution, the system logs the progress and status of each stage.

### 3. Review and Approval (Workflow Approver):

- When a workflow reaches an approval stage, the designated Workflow Approver is notified.
- The Approver accesses a "File Work" module and select type of report in Category option that will display associated file list. To review a file, approver will click on the file name, which open a RHS window, in which Approver navigate to the 'details' section, where Approver must click on the 'Reference File' marked in 'OCR file path', to review that file in the new browsing tab.
- The Workflow Approver reviews the data/output and either approves or rejects it.
- **If Approved:** The workflow proceeds to the next stage.
- **If Rejected:** The workflow either stops, reroutes to a previous stage for correction, or triggers an alert for manual intervention, as defined in the workflow.

### 4. Report Generation:

- Upon successful completion of the workflow (or at a defined reporting stage), the Approver generates reports based on the workflow's outcomes.
- Reports are stored or distributed as configured.

### 5. Monitor Workflow Status: The Clerk (and potentially others) can monitor the status of running and completed workflows through a dashboard, viewing current stage, progress.

#### Postconditions:

- Data is processed automatically according to defined business logic.
- Manual review and approval points ensure data quality and compliance.
- Reports are generated based on processed information.

---

## 3.1.6 Text Analytics (FR-TA-001)

### Use Case: Enhance Data with Text Analytics

**Goal:** To automatically extract meaningful insights (classifications, entities) from text-based data to improve searchability and analysis.

**Actors:** System (automated process, part of FR-WFA-001), Clerk, Data Analyst

#### Preconditions:

- Text-based data (e.g., OCR output, digital document text) has been ingested into the Big Data Repository.
- Text analytics module is configured.

#### Main Flow:

1. **Text Analytics Trigger:** As part of a workflow (FR-WFA-001), after text ingestion or OCR, the workflow triggers the Text Analytics module.
2. **Text Classification (System):**
  - The module processes the ingested text.

- Based on pre-trained models or configurable rules, it assigns categories or tags to the document or specific text segments.
- 3. **Entity Extraction (System):**
  - The module identifies and extracts specific types of entities from the text (e.g., classifications and subtypes, units, locations, events dates).
  - Each extracted entity is categorized by its type.
- 4. **Dictionary Augmentation (System):**
  - The user must import the required data into the dictionary (LHS) to input/edit and append *Location, Orbat and Abbreviation data* that will be access by the system for automatic augmentation as part of Text Analytics configuration.
  - The system will automatically trigger the data augmentation like lat/long, units, formations etc.
- 5. **Store Extracted Information:** The extracted classifications and entities, along with their original text positions, are stored in the Big Data Repository alongside the source document/text.
- 6. **Analytical Querying (Data Analyst):** Data Analysts can now use these classifications and entities as filters or search criteria in the Search Tool (FR-SRCH-001) for more precise analysis.

**Postconditions:**

- Text-based data is enriched with classifications and extracted entities.
  - Data is more structured and discoverable for search and analysis.
- 

### 3.1.7 Geospatial Analysis Module (FR-GIS-001)

#### Use Case: Visualize and Interact with Geospatial Data

**Goal:** To enable users to visualize, filter, and interact with location-based data on an interactive map.

**Actors:** Data Analyst, System User

**Preconditions:**

- Geospatial data (e.g., addresses, GPS coordinates, geocoded locations) is present in the Big Data Repository.
- The GIS module is integrated and operational.

**Main Flow:**

1. **Access GIS Dashboard:** The user navigates to the "Map" module.
2. **Display Base Map:** The system loads a base map (e.g., google map, satellite imagery).
3. **Plot Geospatial Data:** The system automatically plots available geospatial data points (e.g., event locations) from the repository onto the map.

**4. Manage Map Layers:**

- The user can select different map layers (e.g., vector layer, raster layer) to overlay on the base map.
- The user can toggle the visibility of various data sets plotted on the map.

**5. Filter Geospatial Data:**

- The user uses on-map controls or a side panel to filter the displayed data based on attributes (e.g., event date, classification, subtype).
- The map updates on user action to reflect the filtered data.

**6. Interact with Map Elements:**

- The user clicks on a plotted point or area to view detailed information (e.g., event date, classification, subtype).
- The user can select multiple points to perform aggregate analysis.

**7. Plot Custom Elements:**

- The user can manually add custom points, define areas (e.g., polygons for area of interest), or draw routes directly on the map.
- These custom elements can be ingested, saved and exported (in kml, shp and csv formats).

**8. Geo-Fencing and Geo-Search:**

- User can mark the particular area of interest as shapes (like polygon, circle, square) to filter (events, raster, vector layers) and search the data accordingly for specific map type (cluster, pushpin, bubbles).

**9. Geo-Temporal Analysis:** The system should create a timeline layer on top of the geo-layer to determine the temporal-geospatial relationship.

**10. Visualize Reports/Analysis:** The system can overlay visual representations of reports or analytical findings that have a geospatial component (e.g., heatmaps of activity, clusters, pushpins, timeline, synchronisation maps).

**Alternative Flows:**

- **No geospatial data:** The map loads but indicates no plottable data.

**Postconditions:**

- Users have an intuitive visual understanding of location-based data.
- Geospatial data can be analyzed and understood in context.

---

### 3.1.8 Temporal Analysis Module (FR-TIM-001)

#### Use Case: Analyze Events on a Timeline

**Goal:** To provide a chronological visualization of events from diverse data sources, enabling temporal analysis.



**Actors:** Data Analyst, System User

**Preconditions:**

- Data with associated timestamps is present in the Big Data Repository.
- The Temporal Analysis module is operational.

**Main Flow:**

1. **Access Timeline Module:** To access the "Timeline" module, user must perform a Global Search, and in the search result if 'Events' are presents, then open this index by double click on it. Now select the specific number of records and add them to timeline (via right context menu), a new tab with title 'Timeline' will plot all these events in a timeline view.
2. **Plot Events on Timeline:** The user has to input event date, subject, description, and colour & instance (optional). Post that 'Show Timeline' button will be visible to renders the selected events on a chronological timeline, with events marked at their respective timestamps.
3. **Navigate Timeline:**
  - The user can zoom in and out on the timeline to view different time granularities (e.g., hourly, daily, monthly, yearly).
  - The user can pan along the timeline to move through different time periods.
  - The user can filter events by date range or other attributes directly on the timeline.
4. **View Event Details:** When the user clicks or hover on an event marker on the timeline, a detailed tooltip displays information about that specific event (e.g., full text).
5. **Compare Events:** The system allows for viewing multiple types of events simultaneously on the same timeline, enabling comparison and identification of correlations.

**Alternative Flows:**

- **Overlapping events:** The timeline gracefully handles multiple events occurring at the same or very close times (e.g., stacking).

**Postconditions:**

- Users can visually identify patterns, trends, and anomalies over time.
- Temporal relationships between disparate events become apparent.

---

### 3.1.9 Graph Analysis Module (FR-GRPH-001)

**Use Case: Visualize and Explore Data Relationships**

**Goal:** To represent and explore complex relationships between data points (nodes) in an interactive graph visualization.

**Actors:** Data Analyst, System User

**Preconditions:**

- Data with defined relationships (from FR-BDR-001.3 or explicitly defined) exists in the Big Data Repository.
- The Graph Analysis module is operational.

**Main Flow:**

1. **Access Graph Module:** To access the "Graph" module, user must perform a Global Search, and in the search result if 'Events' are presents, then open this index by double click on it. Now select the specific number of records and add them to 'Link-Analysis' (via right context menu), a new tab with title 'Link-Analysis' will plot all these events in a graphical view.
2. **Load Graph:**
  - Automatically: The system plots the graph automatically post user trigger.
  - Explode by Relation Node: User must define the criteria such as relationship type like explode by person, location, unit to plot the graph.
3. **Initial Graph Visualization:** The system renders an interactive graph where data points are represented as nodes and relationships as edges.
4. **Explore Connections:**
  - The user can click on a node to view its immediate connections and associated properties in the RHS panel.
  - The user can expand a node to reveal its connected nodes and their relationships, extending the graph view.
5. **Apply Layouts:** The user can select different graph layout (e.g., spread, circular) to optimize the visual representation for better understanding.
6. **Interact with Graph Elements:** The user can drag nodes, zoom in/out, and pan across the graph to explore different sections.
7. **Save Graph Views:** The user can export specific graph views for future reference or sharing.

**Alternative Flows:**

- **No relationships found:** The system displays individual nodes but indicates no relationships if the selected data is not interconnected.

**Postconditions:**

- Users gain a visual understanding of complex relationships within the data.
- Hidden patterns and connections are revealed through interactive exploration.

---

### 3.1.10 Widgets (FR-WGT-001)

## Use Case: Personalize Dashboards with Widgets

**Goal:** To allow users to create customized dashboards by adding, arranging, and configuring various analytical and informational widgets.

**Actors:** GSO1 (admin role), Col GS-Int (admin role), BGS (admin role).

**Preconditions:**

- The system has available data and analytical outputs.
- User has permission to create and modify dashboards.

**Main Flow:**

1. **Access Dashboard Customization:** The user navigates to their personal dashboard or a shared dashboard they have permission to customize.
2. **Browse Widget Catalogue:** The user clicks an "Add Widget" button, revealing a catalogue of available widgets.
3. **Select and Add Widget:** The user selects a desired widget type (e.g., "Bar Chart", "GIS") or filter it using specified criteria.
4. **Configure Widget:**
  - **For data-driven widgets (charts):** The user selects the data source, metrics, and dimensions to display. They may also configure chart types, colors, and titles.
  - **For informational widgets (maps, cards):** The user may configure the content or source of the information.
5. **Place and Resize Widget:** The newly added widget appears on the dashboard. The user can drag and drop it to rearrange its position and resize it to fit their layout.
6. **Remove Widget:** The user can select an existing widget and choose to remove it from the dashboard.
7. **Rearrange Widgets:** The user can drag and drop widgets to change their order and layout on the dashboard.
8. **Save Dashboard Layout:** The system automatically saves the user's dashboard layout and widget configurations.

**Postconditions:**

- User has a personalized dashboard tailored to their information needs.
- Key information and analysis are readily visible upon logging in.

---

### 3.1.11 Report Generation (FR-RPT-001)

#### Use Case: Generate and Manage Output Reports

**Goal:** To enable users to generate comprehensive reports in various formats, based on defined criteria or workflow outcomes, and manage these reports.

**Actors:** Clerk, GSO1 (admin role), Col GS-Int (admin role), BGS (admin role).

**Preconditions:**

- Data exists in the Big Data Repository.
- Report templates or criteria are defined.
- Workflows that trigger reports are configured (for FR-WFA-001.5).

**Main Flow:**

**1. Select Report Criteria for Workflow Automation:**

- The Clerk accesses the "File Work" interface.
- The Clerk successfully ingest the multiple source files into the system.
- The Clerk will forward the uploaded source file to second higher designation.
- Second designated user chooses relevant data from it, to generate the reports and mark it to the approver.
- The approver may recommend edits, or forward input as is to higher designation.
- Then user will
- The Data Analyst navigates to the "Report Generation" module.
- The Data Analyst selects a pre-defined report type or defines custom criteria for the report (e.g., data filters, specific metrics, time range).

**2. Report Generation for Workflow Automation:**

- Upon successful completion of the workflow (or at a defined reporting stage), the Approver generates reports based on the workflow's outcomes.
- Reports are stored or distributed as configured.

**3. Report Generation from the Search Data:** To generate such report user have to export the grid-data into a desired output format for the report (Word, Excel, PDF, XML, HTML).

- The system processes the data based on the criteria, applies the chosen format, and generates the report file.

**4. Manage Generated Reports through the Workflow Automation:**

- The system stores the generated report in a central repository.
- The user can access a "File Work" module to view a list of all generated reports.
- From this interface, the user can view, download, or archive reports.

**Alternative Flows:**

- **Report generation failure:** If report generation fails (e.g., due to data errors), the system logs the error and notifies the user.
- **Large report size:** The system handles large reports efficiently, potentially offering options for background generation and notification upon completion.

**Postconditions:**

- Reports containing relevant data are generated in desired formats.
- Reports are easily accessible and manageable.

---

### 3.1.12 AI/LLM Model (FR-LLM-001)

#### Use Case: Leverage LLM for Enhanced Search and Summarization

**Goal:** To use a Large Language Model (LLM) to provide more intuitive search and powerful summarization capabilities for ingested data.

**Actors:** Clerk, GSO1 (admin role), Col GS-Int (admin role), BGS (admin role).

**Preconditions:**

- Data is ingested and available in the Big Data Repository (FR-BDR-001).
- The LLM is integrated, trained/fine-tuned, and operational.
- LLM has access to the indexed data (FR-SRCH-001.1).

**Main Flow:**

**1. Enhanced Natural Language Search and Q&A:**

- The user navigates to the search interface to search a keyword of interest, which will render a list of indexes. From these indexes user can navigate to desired data view, where he/she can select the records on which he/she wants to AI/LLM module.
- Here user types a natural language question (e.g., "Give the list of training locations of each unit?").
- The system passes the query to the LLM.
- The LLM processes the query, understands the intent, and retrieves relevant information from the underlying data.
- The LLM synthesizes an answer in natural language, citing sources from the ingested data.
- The answers are displayed to the user and can be copied or exported.

**2. Summarize Events/Documents:**

- The user selects one or more documents, events (from timeline), or profiles (e.g., individual, organization, unit) from the system interface.
- The user initiates a "Summarize" action.
- The system sends the selected content to the LLM.
- The LLM analyzes the content and generates a concise summary.
- The summary is displayed to the user and can be copied or exported.

**3. AI Predictions:**

- The user selects one or more documents, events (from timeline), or profiles (e.g., individual, organization, unit) from the system interface.
- The user initiates an "AI Predictions" action.
- The system sends the selected content to the LLM.
- The LLM analyzes the content and generates a concise prediction.
- The prediction is displayed to the user and can be copied or exported.

#### 4. Knowledge base

- User can access this feature through the right-hand side menu option, named 'AI/LLM', and its submenu having title *Knowledge Base*.
- Here user types a natural language question (e.g., "Give me the name of the permanent member countries of BRICS?").
- The system passes the query to the LLM.
- The LLM processes the query, understands the intent, and retrieves relevant information from the underlying data.
- The LLM synthesizes an answer in natural language, citing sources from the ingested data.
- The answers are displayed to the user and can be copied or exported.

#### 5. Translate/Transliterate:

- User can access this feature through the right-hand side menu option, named 'AI/LLM', and its submenu having title *Translate or Transliterate*.
- Here user can provide the text contents to be translated in desired languages like Mandarin, Urdu, Arabic, Hindi, English, Punjabi.
- The system passes the query to the LLM.
- The LLM processes the text, understands the intent, and retrieves relevant information from the underlying data.
- The LLM synthesizes an answer in translated natural language.
- The translated text will be displayed to the user that can be copied or exported.

#### Alternative Flows:

- **Insufficient information:** If the LLM cannot find enough relevant information to answer a question or summarize, it informs the user.

#### Postconditions:

- Users can interact with the data using natural language.
- Complex information from multiple sources can be quickly summarized.
- Search results are more contextual and direct.

---

### 3.1.13 Predictive Module (FR-PRD-001)

#### Use Case: Define, Configure, and Generate Predictions

**Goal:** To allow a user to access predictive generations and pattern reorganisations based on historical data and generate predictions for various scenarios.

**Actors:** Clerk, GSO1 (admin role), Col GS-Int (admin role), BGS (admin role).

#### Preconditions:

- The user has appropriate permissions to access and configure the Predictive Module.
- Historical data for the past three years (relevant to the chosen use case) is available in the Big Data Repository (FR-BDR-001).

- The system's mathematical modelling and pattern recognition components are operational.

#### **Main Flow:**

- **Access Predictive Module:** The designated user logs into the system and navigates to the "Predictive Alerts" or "Predictive Module" section.
- **Configure Prediction Model:**
  - **Data Selection (FR-PAN-001.4):** The user specifies the relevant historical data sources (from the Big Data Repository) to be used for the prediction, ensuring data spans the last three years.
  - **Pattern Recognition Configuration (FR-PAN-001.2):** The system configures parameters for the pattern recognition algorithms (e.g., temporal windows, locations).
  - **Trend Identification Configuration (FR-PAN-001.5):** The system uses parameters to identify trends from the historical data (e.g., likely hotspots of infiltrations, likely locations of units for training).
  - **Prediction Scenario Mapping (FR-PAN-001.6):** The user links the identified trends and patterns to the specific prediction types (e.g., "infiltration events" lead to "likely hotspots for infiltration").
  - **AI based Model Parameters (FR-PAN-001.7):** The system implicitly uses a mathematical model based on probabilities to deliver emergent patterns, probability scores of events, locations etc.
- **Train/Build Prediction Model (System):** The system processes the selected historical data, identifies patterns and trends, and builds/optimizes the predictive model based on the defined Use Case parameters and mathematical model.
- **Generate Predictions (FR-PAN-001.1, FR-PAN-001.6):**
  - The system triggers the prediction generation for the configured Use Case.
  - The system applies the trained model to current or future-projected data.
  - Predictions are generated for the specific scenarios (e.g., "Likely hotspots for infiltration," "Units which may go for training or move," "Likely hotspots (concentration) of radars").
- **View and Analyze Predictions:**
  - The system displays the generated predictions, potentially on relevant dashboards (e.g., GIS dashboard for hotspots, timeline for unit movements).

#### **Postconditions:**

- Predictions for specific scenarios are generated and available for viewing and further analysis.
- The system has performed pattern recognition and trend identification based on historical data.

---

### 3.1.14 IMINT Module (FR-IMT-001)

#### Use Case: Ingest IMINT Data into Form-Based Collation System

**Goal:** To enable users to ingest Imagery Intelligence (IMINT) data from various sources into a standardized, form-based collation system.

**Actors:** Clerk, GSO1 (admin role), Col GS-Int (admin role), BGS (admin role).

**Preconditions:**

- The designated user has appropriate permissions to ingest data.
- The system's form-based collation system is operational.
- For trial purposes, the IMINT data in its existing format is accessible.
- For operational ingestion, IMINT from the IFC workstation of IIT is accessible.

**Main Flow:**

1. The user navigates to the "File Work" module and click on the 'Create' button for data ingestion. Add the mandatory fields like file, name, report date, file type for the datasets (file) you wish to ingest.
2. **Configure Pipeline:**
  - **For File-based sources:** The user specifies the file path(s) or directory, file naming conventions.
  - **Select Source Type:** The user chooses the data source type from a list:
    - Scanned Documents (FR-ING-001.4)
    - Digital Document Files (e.g., PDF, Word, PPTX) (FR-ING-001.5)
3. **Execute Ingestion:**
  - The user clicks "submit" to initiate the ingestion process immediately.
  - The system implicitly maps the ingested data to the Big Data Repository.
4. **Ingestion Status:** The system displays a fadeout message with ingestion results.
5. **Address Errors (if any):** If failures occur, the user can review the logs, identify the root cause, and take corrective actions (e.g., correct source data) and re-run the ingestion.

**Postconditions:**

- Relevant IMINT data is successfully ingested and stored within the form-based collation system.
  - The IMINT Ingestion Dashboard provides visibility into recent ingestion activities.
- 
-



## Appendix B: Data Dictionary

- **Case Table**
    - CaseID: Integer, Primary Key, Auto-increment
    - CaseTitle: Varchar(255), Not Null
    - CaseDescription: Text, Nullable
    - Priority: Varchar(10), Not Null (High, Medium, Low)
    - AssignedTo: Integer, Foreign Key (UserID), Nullable
    - StartDate: Date, Not Null
    - CreatedDate: DateTime, Not Null
    - LastModifiedDate: DateTime, Nullable
  - **User Table**
    - UserID: Integer, Primary Key, Auto-increment
    - Username: Varchar(50), Unique, Not Null
    - PasswordHash: Varchar(255), Not Null
    - FirstName: Varchar(50), Not Null
    - LastName: Varchar(50), Not Null
    - Role: Varchar(50), Not Null
  - **SearchIndex Table**
    - IndexID: Integer, Primary Key, Auto-increment.
    - IndexName: Varchar(100), Unique, Not Null
    - IndexDescription: Text, Nullable
    - IndexLocation: Varchar(255), Not Null
  - **Document Table**
    - DocumentID: Integer, Primary Key, Auto-increment
    - CaseID: Integer, Foreign Key (CaseID), Nullable
    - DocumentName: Varchar(255), Not Null
    - DocumentPath: Varchar(255), Not Null
    - DocumentType: Varchar(50), Not Null
    - UploadDate: DateTime, Not Null
- 

## Appendix C: Glossary

- **Data Ingestion:** The process of importing data from various sources into the system.
  - **Data Normalization:** The process of organizing data to reduce redundancy and improve consistency.
  - **Fuzzy Search:** A search that finds results even when the search term is misspelled or contains variations.
  - **Knowledge Graph:** A graph-based representation of entities and their relationships.
  - **RBAC (Role-Based Access Control):** A method of regulating access to computer or network resources based on the roles of individual users within an organization.
  - **SSO (Single Sign-On):** An authentication process that allows users to access multiple applications with a single set of login credentials.
-

## Appendix D: Acceptance Criteria

- **FR1: Data Ingestion**
    - Given a CSV file, when the user uploads it, then the system shall successfully import the data into the database.
    - Given a JSON API endpoint, when the system polls the endpoint, then the system shall successfully import the data.
  - **FR12: Keyword Search**
    - Given multiple indexes containing data, when a keyword is entered, then the system shall display results from all relevant indexes.
    - Given a large dataset, when a keyword search is performed, then results will be displayed in less than 3 seconds.
  - **NFR7: Security**
    - When a SQL injection attack is attempted, then the system shall prevent the attack and log the attempt.
    - When an XSS attack is attempted, then the system shall sanitize the input and prevent the attack.
- 

## Appendix E: Risk Analysis

- **Risk 1: Data Security Breach**
    - **Likelihood:** Medium
    - **Impact:** High
    - **Mitigation:** Implement strong encryption, access controls, and regular security audits.
  - **Risk 2: Performance Degradation**
    - **Likelihood:** Medium
    - **Impact:** Medium
    - **Mitigation:** Implement load balancing, database optimization, and performance monitoring.
  - **Risk 3: Data Ingestion Failure**
    - **Likelihood:** Low
    - **Impact:** Medium
    - **Mitigation:** Implement robust error handling, data validation, and retry mechanisms.
- 

## Appendix F: Traceability Matrix

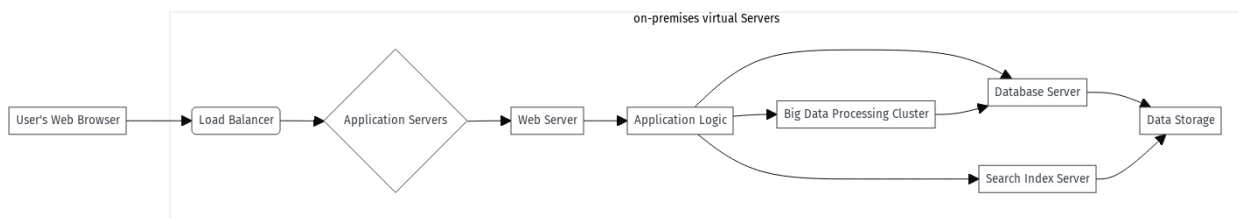
This matrix maps requirements to design elements, code modules, and test cases, ensuring that all requirements are addressed.

Requirement ID	Design Element	Code Module	Test Case ID	Verification Method
FR1: Data Ingestion	DataIngestionService	DataIngestion.py	TC-FR1-1, TC-FR1-2	Integration Test
FR12: Keyword Search	SearchService	Search.py	TC-FR12-1, TC-FR12-2	Integration Test
NFR7: Security (SQL Injection)	SecurityFilter	Security.py	TC-NFR7-1	Security Scan, Penetration Test
UI1: Web Interface	UserInterface	ui/index.html	TC-UI1-1	Manual Test, UI Automation
UC1: Create New Case	CaseController	cases/create.py	TC-UC1-1, TC-UC1-2	Manual Test, Integration test
NFR17: Reliability (24/7)	LoadBalancer, DatabaseReplication	infrastructure/	TC-NFR17-1	Stress test, failover test

## Appendix G: Business Rules

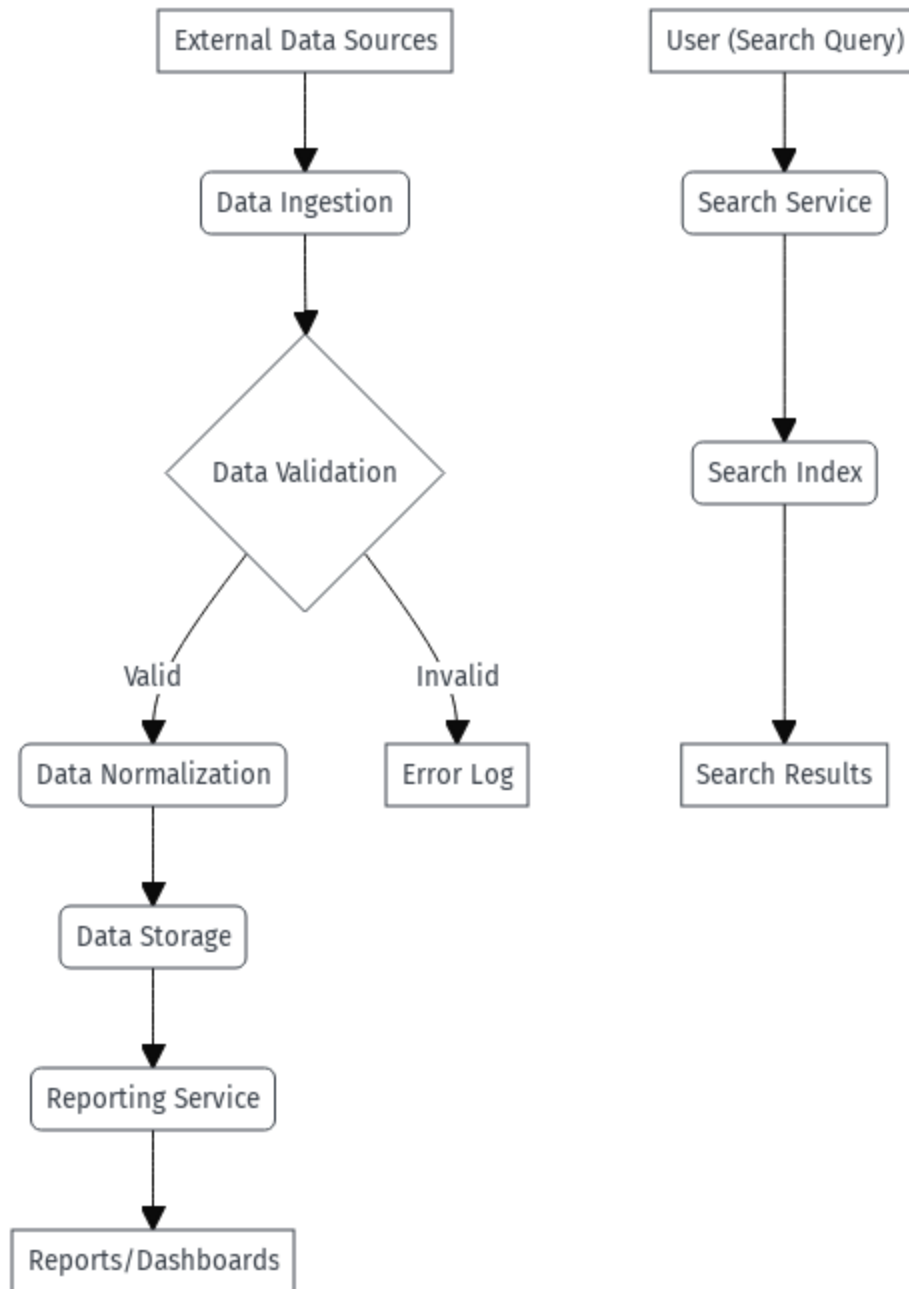
- **BR1:** All cases must have a unique case ID.
- **BR2:** User roles must be assigned according to the organization's security policy.
- **BR3:** Data retention policies must comply with legal requirements.
- **BR4:** Search results must be ranked by relevance.
- **BR5:** Only authorized users can access sensitive data.
- **BR6:** Audit logs must be maintained for all system activities.

## Appendix H: Deployment Diagram



- **Description:** This diagram shows the physical deployment of the system, including the load balancer, application servers, database server, search index server, and big data processing cluster.

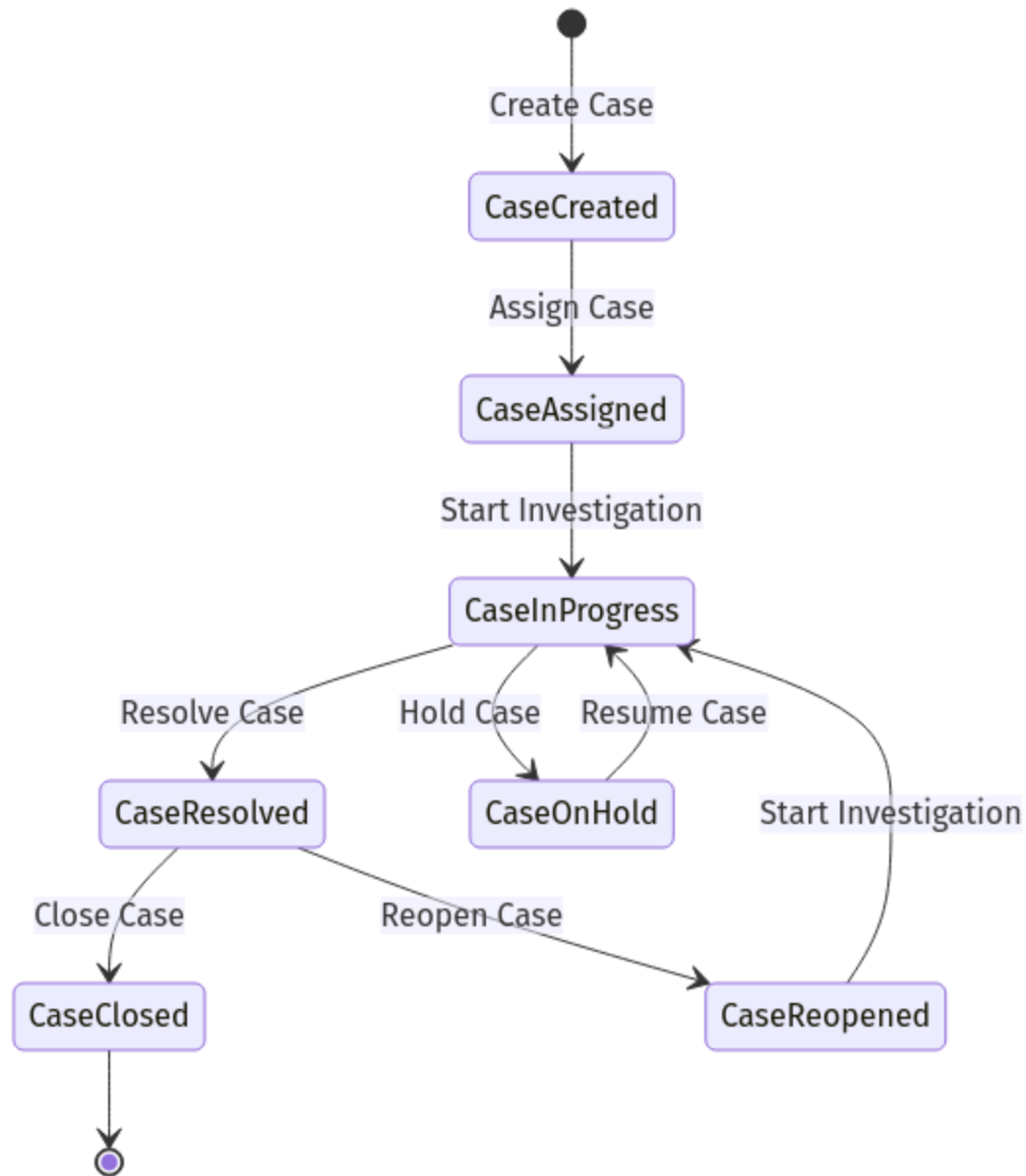
## Appendix I: Data Flow Diagram



- **Description:** This diagram illustrates the flow of data through the system, from external data sources to data storage and reporting.

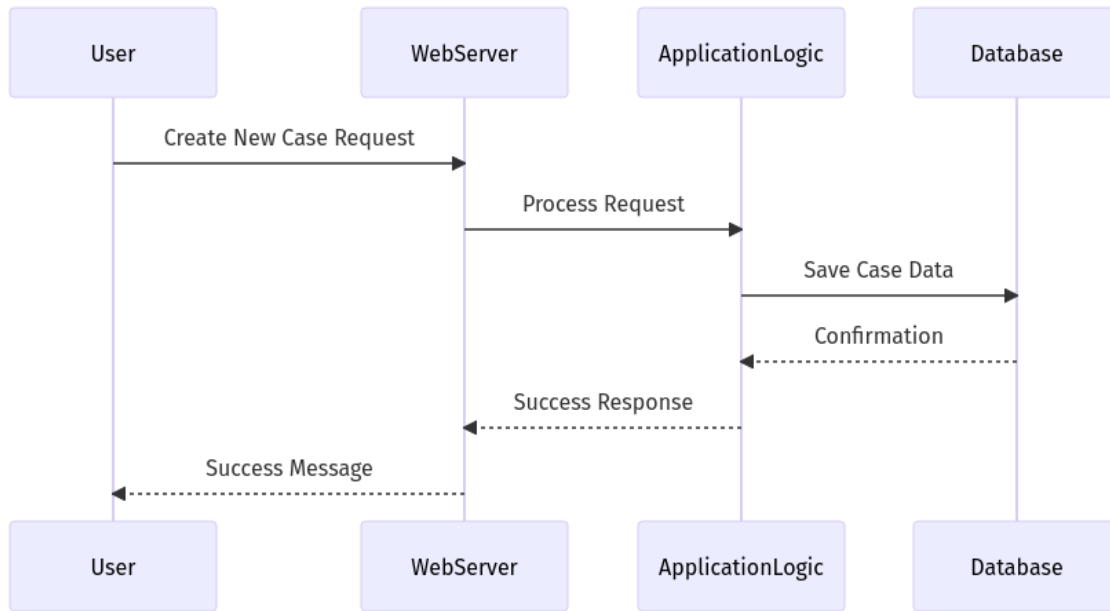
---

## Appendix J: State Transition Diagram



- **Description:** This diagram shows the possible states of a case and the transitions between them.

## Appendix K: Sequence Diagram



- **Description:** This diagram illustrates the sequence of interactions between the user, web server, application logic, and database when creating a new case.

## Appendix L: Test Cases

### Test Cases for Configure and Execute On-Demand Data Ingestion

Test Case ID	Description	Preconditions	Steps	Expected Result
TC-ING-001	Successful Excel File Ingestion	User (Clerk) with permissions; Excel file valid_data.xlsx (e.g., 5 rows, 3 columns) accessible.	1. User navigates to "File Work" module, clicks "Create." 2. Fills mandatory fields: file_name, report_date, file_type=Excel. 3. Specifies valid_data.xlsx path. 4. Clicks "submit."	Fadeout message: "Ingestion successful." Data from valid_data.xlsx (5 records) is in Big Data Repository.
TC-ING-002	Successful Access Database Ingestion	User (Clerk) with permissions; Access DB sample.accdb with table Customers (e.g., 10 records) accessible.	1. User navigates to "File Work" module, clicks "Create." 2. Fills mandatory fields: database_name, report_date, file_type=Access. 3. Provides connection details for sample.accdb, selects Customers table. 4. Clicks "submit."	Fadeout message: "Ingestion successful." Data from Customers table (10 records) is in Big Data Repository.
TC-ING-	Successful SQL	User (Clerk) with permissions; SQL	1. User navigates to "File Work" module, clicks "Create." 2. Fills	Fadeout message: "Ingestion

003	<b>Database Ingestion</b>	DB prod_db with table Orders (e.g., 20 records) accessible.	mandatory fields: database_name, report_date, file_type=SQL. 3. Provides connection details for prod_db, selects Orders table. 4. Clicks "submit."	successful." Data from Orders table (20 records) is in Big Data Repository.
TC-ING-004	<b>Successful Scanned Document Ingestion (PNG)</b>	User (Clerk) with permissions; invoice.png (clear scanned image) in a specified directory accessible.	1. User navigates to "File Work" module, clicks "Create." 2. Fills mandatory fields: document_name, report_date, file_type=Scanned Document. 3. Specifies invoice.png path. 4. Clicks "submit."	Fadeout message: "Ingestion successful." invoice.png is in Big Data Repository. (OCR will be triggered later).
TC-ING-005	<b>Successful Digital Document Ingestion (PDF)</b>	User (Clerk) with permissions; report.pdf (digital PDF) in a specified directory accessible.	1. User navigates to "File Work" module, clicks "Create." 2. Fills mandatory fields: document_name, report_date, file_type=Digital Document. 3. Specifies report.pdf path. 4. Clicks "submit."	Fadeout message: "Ingestion successful." report.pdf is in Big Data Repository with text extracted.
TC-ING-006	<b>Ingestion Failure - Non-existent File Path</b>	User (Clerk) with permissions.	1. User navigates to "File Work" module, clicks "Create." 2. Fills mandatory fields with valid data. 3. Specifies non-existent file path (e.g., C:\non\existent\file.xlsx). 4. Clicks "submit."	Fadeout message indicates ingestion failure (e.g., "File Not Found"). Error log available with details.
TC-ING-007	<b>Ingestion Failure - Invalid Database Credentials</b>	User (Clerk) with permissions; SQL DB secure_db requires valid credentials.	1. User navigates to "File Work" module, clicks "Create." 2. Fills mandatory fields with valid data. 3. Provides connection details for secure_db with incorrect credentials. 4. Clicks "submit."	Fadeout message indicates ingestion failure (e.g., "Authentication Failed"). Error log available with details.
TC-ING-008	<b>Ingestion Failure - Data Format Issues in Excel</b>	User (Clerk) with permissions; Excel file bad_format.xlsx with a column expecting numbers but containing text.	1. User navigates to "File Work" module, clicks "Create." 2. Fills mandatory fields. 3. Specifies bad_format.xlsx path. 4. Clicks "submit."	Fadeout message indicates partial/full ingestion failure (e.g., "Data Parsing Error"). Error log available detailing specific row/column errors. Some valid records may still be ingested.
TC-ING-009	<b>Re-running Ingestion After Error Correction</b>	User (Clerk) with permissions; Scenario where TC-ING-006 (non-	1. User reviews logs from failed TC-ING-006. 2. Corrects the file path to C:\actual\existing\file.xlsx in the	Fadeout message: "Ingestion successful." Data from file.xlsx is in

		existent file) has occurred.	pipeline configuration. 3. Re-runs the ingestion pipeline.	Big Data Repository.
--	--	------------------------------	--	----------------------

## Test Cases for Perform Multilingual OCR on Scanned Documents

Test Case ID	Description	Preconditions	Steps	Expected Result
TC-OCR-001	Automatic OCR Trigger on English Document Ingestion	english_scan.png (clear, English text) ingested via FR-ING-001.4.	1. Ingest english_scan.png. 2. System automatically processes OCR. 3. Verify extracted text in the document's detail view.	Extracted text is present and accurately reflects the English content of the image.
TC-OCR-002	Automatic OCR Trigger on Hindi Document Ingestion	hindi_scan.png (clear, Hindi text) ingested via FR-ING-001.4.	1. Ingest hindi_scan.png. 2. System automatically processes OCR. 3. Verify extracted text in the document's detail view.	Extracted text is present and accurately reflects the Hindi content of the image.
TC-OCR-003	Automatic OCR Trigger on Mandarin Document Ingestion	mandarin_scan.png (clear, Mandarin text) ingested via FR-ING-001.4.	1. Ingest mandarin_scan.png. 2. System automatically processes OCR. 3. Verify extracted text in the document's detail view.	Extracted text is present and accurately reflects the Mandarin content of the image.
TC-OCR-004	OCR Error Review and Correction (English)	english_scan_with_error.png (contains "receipt" instead of "receipt") ingested, OCR outputs "receipt".	1. Data Reviewer accesses the document in the review interface. 2. Observes "receipt" in extracted text. 3. Edits the text to "receipt". 4. Saves changes.	The extracted text for english_scan_with_error.png is updated to "receipt" in the Big Data Repository.



<b>TC-OCR-005</b>	<b>OCR Error Review and Correction (Hindi - Language Detection Issue)</b>	hindi_scan_blur.png (Hindi text, slightly blurry, OCR incorrectly detects as English or mixes languages leading to poor extraction).	1. Ingest hindi_scan_blur.png. 2. Data Reviewer accesses the document in the review interface. 3. Observes poor or incorrect extracted text. 4. (If applicable) Attempts to manually override the detected language to Hindi (if the UI allows, otherwise proceeds with manual correction). 5. Manually corrects significant portions of the text. 6. Saves changes.	The extracted text for hindi_scan_blur.png is updated to the manually corrected Hindi text in the Big Data Repository. The system should ideally flag potential language detection issues.
<b>TC-OCR-006</b>	<b>OCR Performance on Mixed Language Document</b>	mixed_lang_doc.png (contains blocks of English, Hindi, and Mandarin text).	1. Ingest mixed_lang_doc.png. 2. System automatically processes OCR. 3. Verify extracted text for all three languages.	All three languages' texts are extracted, though accuracy might vary. The extracted text should be a combination of recognized English, Hindi, and Mandarin.
<b>TC-OCR-007</b>	<b>No OCR Performed for Non-Image Files</b>	word_doc.docx (digital Word document) ingested via FR-ING-001.5.	1. Ingest word_doc.docx. 2. Verify document's content.	Text is extracted directly from the DOCX (not via OCR). No OCR process is triggered for this file type.

## Test Cases for Manage and Query Unified Data Repository

Test Case ID	Description	Preconditions	Steps	Expected Result
<b>TC-BDR-001</b>	<b>Successful Data Compilation from Multiple Sources</b>	TC-ING-001 (Excel), TC-ING-003 (SQL), TC-OCR-001 (Scanned+OCR) have been executed successfully.	1. Verify all ingested data types (Excel records, SQL records, OCR text from image) are visible in the Big Data Repository. 2. Perform a basic search	All data from the respective ingestion pipelines is present and searchable in the Big Data Repository.

			across all data types (e.g., search for a common keyword from Excel, a record ID from SQL, and a word from OCR text).	
<b>TC-BDR-002</b>	<b>Automatic Data Co-relation (Event ID)</b>	excel_events.xlsx (contains Event_ID, Event_Desc) ingested. sql_incidents.db (contains Incident_ID, Event_ID, Incident_Date) ingested. Co-relation rule defined for Event_ID.	1. Ingest excel_events.xlsx. 2. Ingest sql_incidents.db. 3. Perform a query/search using Event_ID.	Query results show linked data from both Excel and SQL sources based on Event_ID (e.g., Event_Desc from Excel displayed alongside Incident_Date from SQL for the same Event_ID).
<b>TC-BDR-003</b>	<b>Handling Structured Data Storage</b>	TC-ING-003 (SQL) executed successfully.	1. Access the stored data for Orders table in the Big Data Repository. 2. Verify schema and data types.	The Orders data retains its tabular structure, column names, and original data types (e.g., dates as dates, numbers as numbers).
<b>TC-BDR-004</b>	<b>Handling Unstructured Data Storage</b>	TC-OCR-001 (Scanned+OCR) executed successfully.	1. Access the english_scan.png entry in the Big Data Repository. 2. Verify the associated extracted text.	The extracted text from english_scan.png is stored as searchable unstructured text, linked to the original image.
<b>TC-BDR-005</b>	<b>Query Execution Error - Malformed Query</b>	User (Data Analyst) with query permissions.	1. User attempts to execute a custom query with a syntax error (e.g., SELECT FROM Customers).	System displays an informative error message indicating a malformed query syntax. No results are returned.
<b>TC-BDR-006</b>	<b>Query Execution Error - Non-existent Data Target</b>	User (Data Analyst) with query permissions.	1. User attempts to execute a custom query targeting a non-existent table/data type (e.g., SELECT * FROM NonExistentTable).	System displays an informative error message indicating the target data does not exist. No results are returned.

## Test Cases for Search and Retrieve Information

Test Case ID	Description	Preconditions	Steps	Expected Result
TC-SRCH-001	<b>Global Search (360 Search) - Single Keyword</b>	Big Data Repository contains various documents/records with the keyword "financial."	1. System User navigates to "Search" module. 2. Enters "financial" in the global search bar. 3. Clicks "Search."	Tree-map displays results from all indexes containing "financial," with volume indicators. Results list shows relevant documents/records from various sources (e.g., Excel, PDF, SQL) containing "financial."
TC-SRCH-002	<b>Global Search (360 Search) - Group of Keywords</b>	Big Data Repository contains documents/records with "project alpha" and "budget."	1. System User navigates to "Search" module. 2. Enters "project alpha budget" in the global search bar. 3. Clicks "Search."	Tree-map displays results related to "project alpha" and "budget." Results list shows documents/records where both keywords appear, prioritized by relevance.
TC-SRCH-003	<b>Basic Search - Single Parameter (Document Type)</b>	Repository contains "Report" and "Contract" document types.	1. System User navigates to "Search" module. 2. Selects "Document" from the structured fields dropdown/input. 3. Enters "Report" as the criteria. 4. Clicks "Search."	Results tree-map and list exclusively display documents categorized as "Report," indicating their source.
TC-SRCH-004	<b>Basic Search - Multiple Parameters (Event Type &amp; Date)</b>	Repository contains "Meeting" and "Incident" events across various dates.	1. System User navigates to "Search" module. 2. Selects "events" from structured fields, enters "Meeting." 3. Selects "report date" field, enters a specific date range (e.g., "01/01/2025 to 01/31/2025"). 4. Clicks "Search."	Results tree-map and list display only "Meeting" events that occurred within January 2025.

TC-SRCH-005	Advanced Search - Boolean AND	Repository contains documents with "acquisition" and others with "merger"; some contain both.	1. Data Analyst navigates to "Search" module, selects "Advanced Search." 2. Enters "acquisition AND merger." 3. Clicks "Search."	Results tree-map and list display only documents/records containing <i>both</i> "acquisition" and "merger."
TC-SRCH-006	Advanced Search - Boolean OR	Repository contains documents with "fraud" and others with "embezzlement."	1. Data Analyst navigates to "Search" module, selects "Advanced Search." 2. Enters "fraud OR embezzlement." 3. Clicks "Search."	Results tree-map and list display documents/records containing <i>either</i> "fraud" <i>or</i> "embezzlement" (or both).
TC-SRCH-007	Advanced Search - Boolean NOT	Repository contains documents about "project X" and some about "project X" that mention "delay."	1. Data Analyst navigates to "Search" module, selects "Advanced Search." 2. Enters "project X NOT delay." 3. Clicks "Search."	Results tree-map and list display documents/records containing "project X" but <i>excluding</i> those that also mention "delay."
TC-SRCH-008	Advanced Search - Proximity Search	Repository contains text where "security" and "breach" appear close together and far apart.	1. Data Analyst navigates to "Search" module, selects "Advanced Search." 2. Enters "security /5 breach" (assuming /5 is the syntax for 5 words proximity). 3. Clicks "Search."	Results tree-map and list display documents/records where "security" and "breach" are within 5 words of each other, prioritizing closer proximity.
TC-SRCH-009	Advanced Search - Fuzzy Search	Repository contains "receipt," "receiept," "recept."	1. Data Analyst navigates to "Search" module, selects "Advanced Search." 2. Enters "receipt~" (assuming ~ is fuzzy indicator). 3. Clicks "Search."	Results tree-map and list display documents/records containing "receipt," "receiept," and "recept."
TC-SRCH-010	Refine Results - Filter by Source	Initial search returns results from Excel, PDF, and SQL sources.	1. Perform a global search (e.g., "report"). 2. From the filter options, select "Source: PDF."	The displayed results tree-map and list update to show only PDF documents matching "report."
TC-SRCH-011	Sort Results - By Date (Newest)	Initial search returns results from various dates.	1. Perform a global search (e.g., "activity log"). 2. Select "Sort by: Date (Newest First)."	The displayed results list reorders to show the most recent

	First)			documents/records first.
TC-SRCH-012	No Results Found	Search for a unique, non-existent keyword (e.g., "xyz_nonexistent_data").	1. System User navigates to "Search" module. 2. Enters "xyz_nonexistent_data" in global search. 3. Clicks "Search."	System displays a message: "No results found matching your criteria." Tree-map is empty.

## Test Cases for Define and Execute Automated Workflows

Test Case ID	Description	Preconditions	Steps	Expected Result
TC-WFA-001	Clerk Ingests Multiple Files & Forwards for Workflow	Clerk has permission; file1.pdf, file2.xlsx, file3.docx are ready for ingestion.	1. Clerk accesses "File Work" interface. 2. Clerk uses ingestion module to successfully ingest file1.pdf, file2.xlsx, file3.docx. 3. Clerk selects ingested files and initiates "Forward to Second Higher Designation" action.	All three files are successfully ingested into the system. The files are transferred to the queue/inbox of the "Second Higher Designation" user.
TC-WFA-002	Second Designated User Chooses Data & Forwards for Approval	file1.pdf (containing relevant data for a report) is in "Second Higher Designation" user's inbox.	1. Second designated user logs in and accesses their workflow inbox. 2. Selects file1.pdf. 3. Extracts/chooses relevant data points from file1.pdf (e.g., by highlighting, marking). 4. Initiates "Generate Report" action based on chosen data. 5. Marks/forwards the report to the "Approver."	A draft report is generated based on the selected data. The report/task is moved to the "Approver's" inbox/queue for review.
TC-WFA-	Approver Reviews and	A report/task (from TC-WFA-	1. Approver logs in, navigates to "File	The Approver can seamlessly review the reference file. The

<b>003</b>	<b>Recommends Edits</b>	002) is in the "Approver's" inbox. file1.pdf is the reference file.	Work" module. 2. Selects "Report" category, finds and clicks on the file name. 3. In the RHS window, navigates to "details" section. 4. Clicks "Reference File" (marked in "OCR file path") to open file1.pdf in a new tab. 5. Reviews the report content against file1.pdf. 6. Enters comments for edits (e.g., "Incorrect figure on page 3") and selects "Recommend Edits" or similar action.	report/task is returned to the "Second Higher Designation" user with the Approver's comments/recommendations for edits.
<b>TC-WFA-004</b>	<b>Approver Approves and Forwards to Higher Designation</b>	A report/task (that may have been edited based on TC-WFA-003, or a new one) is in the "Approver's" inbox, ready for final approval.	1. Approver logs in and reviews the report/task. 2. Clicks "Approve" (or "Forward as is to Higher Designation").	The report/task is marked as "Approved" by the Approver. The report/task is automatically moved to the inbox/queue of the "Higher Designation" user.
<b>TC-WFA-005</b>	<b>Workflow Rejection - Reroute to Previous Stage</b>	A workflow stage is configured to reroute to a previous stage upon rejection. Approver rejects a task.	1. Approver reviews a task/report. 2. Selects "Reject" and provides rejection reason.	The workflow status changes to "Rejected" for that stage. The task is automatically moved back to the designated previous stage (e.g., "Second Higher Designation") with the rejection reason.
<b>TC-WFA-006</b>	<b>Report Generation upon Workflow Completion</b>	A workflow is configured to generate a report upon final approval. The workflow completes.	1. Final Approver approves the workflow (TC-WFA-004). 2. Verify report presence.	A final report is automatically generated based on the workflow's outcomes and stored in the report management system.

## Test Cases for Enhance Data with Text Analytics

Test Case ID	Description	Preconditions	Steps	Expected Result
TC-TA-001	<b>Automatic Text Classification</b>	Ingested report_daily.pdf (digital document with daily-related text). Text analytics module configured for "Daily," "Operations" classifications.	1. Ingest report_finance.pdf (triggers workflow). 2. System triggers Text Analytics module. 3. View extracted classifications for report_daily.pdf.	report_daily.pdf is automatically classified as "DIB" (or a relevant daily sub-category).
TC-TA-002	<b>Entity Extraction - Locations</b>	Ingested events.docx (digital document mentioning "New Delhi," "Mumbai," "Bangalore"). Text analytics module configured for location entity extraction.	1. Ingest events.docx (triggers workflow). 2. System triggers Text Analytics module. 3. View extracted entities for events.docx.	"New Delhi," "Mumbai," "Bangalore" are extracted and categorized as "Location" entities.
TC-TA-003	<b>Entity Extraction - Units</b>	Ingested military_dispatch.txt (raw text mentioning "Alpha Company," "7th Battalion"). Text analytics module configured for unit entity extraction.	1. Ingest military_dispatch.txt (triggers workflow). 2. System triggers Text Analytics module. 3. View extracted entities for military_dispatch.txt.	"Alpha Company," "7th Battalion" are extracted and categorized as "Unit" entities.
TC-TA-004	<b>Dictionary Augmentation - Location (Lat/Long)</b>	User imports a dictionary entry: "Headquarters" -> "12.9716 N, 77.5946 E". Ingested operation_plan.pdf (mentions "Headquarters").	1. User imports the Location dictionary data. 2. Ingest operation_plan.pdf. 3. View extracted entities for operation_plan.pdf.	"Headquarters" is extracted as a "Location" entity, and its corresponding Lat/Long (12.9716 N, 77.5946 E) is automatically augmented/associated with it.
TC-TA-005	<b>Dictionary Augmentation - Abbreviation Expansion</b>	User imports dictionary entry: "IAF" -> "Indian Air Force." Ingested defense_report.docx (mentions "IAF operations").	1. User imports the Abbreviation dictionary data. 2. Ingest defense_report.docx. 3. Perform a search for "Indian Air Force" (via FR-SRCH-001).	defense_report.docx is returned in search results for "Indian Air Force," demonstrating that "IAF" was augmented/expanded.

TC-TA-006	Search Using Extracted Classifications/Entities	TC-TA-001 (classified as "Finance") and TC-TA-002 (locations extracted) are prerequisites.	1. Data Analyst navigates to "Search" module (FR-SRCH-001). 2. Filters search by "Classification: Finance." 3. Performs another search, filtering by "Location: New Delhi."	1. Only documents classified as "Finance" are displayed. 2. Only documents mentioning "New Delhi" as a location are displayed.
TC-TA-007	No Relevant Classifications/Entities Found	Ingested random_notes.txt (contains text not covered by trained models or dictionary entries).	1. Ingest random_notes.txt. 2. System triggers Text Analytics module. 3. View extracted information for random_notes.txt.	No specific classifications or entities are extracted, or only very generic ones if a fallback mechanism exists.

## Test Cases for Visualize and Interact with Geospatial Data

Test Case ID	Description	Preconditions	Steps	Expected Result
TC-GIS-001	Display Base Map & Plot Geospatial Data	Big Data Repository contains 10 event records with valid Lat/Long coordinates.	1. User (Data Analyst) navigates to "Map" module. 2. System loads a base map (e.g., Google Maps imagery). 3. All 10 event locations are automatically plotted as markers on the map.	A functional base map is displayed, and all 10 event data points are visibly plotted on their correct geographical locations.
TC-GIS-002	Manage Map Layers - Toggle Visibility	Map is displaying "Event Locations" and "Unit Deployments" as separate layers.	1. User accesses "Map Layers" control. 2. Toggles "Unit Deployments" layer off. 3. Toggles "Unit Deployments" layer on again.	1. "Unit Deployments" data points disappear from the map while "Event Locations" remain. 2. "Unit Deployments" data points reappear on the map.
TC-	Filter Geospatial	Map displays	1. User accesses the	Only events that



GIS-003	<b>Data - By Event Date</b>	events from various dates.	map filter panel. 2. Selects "Event Date" filter and sets a range (e.g., "last month"). 3. Observes map update.	occurred within the "last month" are displayed on the map; others are hidden.
TC-GIS-004	<b>Interact with Map Elements - View Details</b>	Map displays multiple plotted event points.	1. User clicks on a specific event point on the map.	A detailed information panel/tooltip appears, showing event date, classification, subtype, and other relevant attributes for that specific point.
TC-GIS-005	<b>Plot Custom Element - Polygon Area of Interest</b>	User has permissions to add custom elements.	1. User selects "Draw Polygon" tool on the map. 2. Draws a polygon defining a specific area (e.g., a city district). 3. Saves the polygon.	A persistent polygon shape is drawn on the map. The system allows naming/saving this custom polygon.
TC-GIS-006	<b>Export Custom Element (KML, SHP, CSV)</b>	A custom polygon (from TC-GIS-005) is saved.	1. User selects the saved custom polygon. 2. Selects "Export" option. 3. Chooses "KML" format and downloads. 4. Repeats for "SHP" and "CSV" formats.	The custom polygon data is successfully downloaded in KML, SHP, and CSV formats, containing appropriate geospatial data.
TC-GIS-007	<b>Geo-Fencing/Geo-Search - Polygon Filter (Events)</b>	Map displays events. A custom polygon (from TC-GIS-005) is drawn.	1. User selects the drawn polygon. 2. Applies "Filter Events within Polygon" action. 3. Selects a map type (e.g., "cluster").	Only event data points falling <i>within</i> the defined polygon are displayed on the map, clustered. Events outside are hidden.
TC-GIS-008	<b>Geo-Temporal Analysis - Timeline Layer</b>	Map displays event data with timestamps.	1. User accesses the "Map" module. 2. Activates the "Timeline Layer" option. 3. Interacts with the timeline (e.g., pans, selects date range).	A timeline appears on top of the map. As the user navigates the timeline, events on the map dynamically appear/disappear based on their temporal relationship to the selected time window.
TC-GIS-009	<b>Visualize Reports/Analysis - Heatmap</b>	Geospatial data with varying densities (e.g.,	1. User selects a report visualization option (e.g.,	A heatmap overlay is displayed on the map, visually representing

		incident frequency in different areas).	"Generate Heatmap"). 2. Specifies data for heatmap.	areas of high/low density (e.g., "hotspots" of incidents).
<b>TC-GIS-010</b>	<b>No Geospatial Data Available</b>	Big Data Repository contains no data with valid geospatial coordinates.	1. User navigates to "Map" module.	A base map loads, but a clear message is displayed (e.g., "No plottable geospatial data found") and no data points appear.

## Test Cases for Analyse Events on a Timeline

Test Case ID	Description	Preconditions	Steps	Expected Result
<b>TC-TIM-001</b>	<b>Access Timeline via Global Search &amp; Event Index</b>	Big Data Repository has "Event" records with event_date, subject, description.	1. User performs a Global Search for "Operation Thunder." 2. In search results, double-clicks the "Events" index. 3. Selects 5 relevant event records (via right-click context menu). 4. Clicks "Add to Timeline" (via right-click context menu).	A new browser tab titled "Timeline" opens, displaying the 5 selected events plotted on a chronological timeline view.
<b>TC-TIM-002</b>	<b>Plot Events on Timeline (Manual Input)</b>	User has access to the "Timeline" module.	1. User navigates to "Timeline" module (or directly opens it if applicable). 2. Enters event_date, subject, description for 3 events. 3. (Optional) Assigns different colors/instances. 4. Clicks "Show Timeline."	The three manually entered events are correctly rendered on the chronological timeline, with specified colors/instances.
<b>TC-TIM-003</b>	<b>Navigate Timeline - Zoom (Yearly to Daily)</b>	Timeline (from TC-TIM-001 or TC-TIM-002) is displayed.	1. User is viewing timeline at "Yearly" granularity. 2. Zooms in progressively (e.g., clicks zoom button, uses scroll wheel).	The timeline granularity changes from "Yearly" to "Monthly," then "Daily," showing events with increasing detail.
<b>TC-TIM-004</b>	<b>Navigate Timeline - Pan</b>	Timeline is displayed.	1. User pans the timeline (drags horizontally).	The timeline view shifts to display different chronological periods.
<b>TC-TIM-</b>	<b>Filter Events by Date</b>	Timeline displays events across	1. User accesses timeline filter controls. 2. Sets a	Only events falling within June 2024 are

005	Range on Timeline	several years.	date range filter (e.g., "2024-06-01" to "2024-06-30").	visible on the timeline.
TC-TIM-006	View Event Details on Hover	Timeline displays multiple event markers.	1. User hovers the mouse pointer over a specific event marker.	A detailed tooltip appears next to the marker, displaying the event's subject, description, and event_date.
TC-TIM-007	Compare Multiple Event Types	Timeline displays "Incident" events and "Training" events.	1. User enables viewing both "Incident" and "Training" event types on the same timeline (e.g., different colors/icons).	Both event types are clearly visible on the same timeline, allowing visual comparison of their occurrences.
TC-TIM-008	Handling Overlapping Events	Two events (Event A, Event B) occur on the exact same date and time.	1. Plot Event A and Event B on the timeline.	The timeline interface displays both events clearly, potentially by stacking them, using offset markers, or providing a mechanism to cycle through them at the same timestamp.

## Test Cases for Visualize and Explore Data Relationships

Test Case ID	Description	Preconditions	Steps	Expected Result
TC-GRPH-001	Access Graph via Global Search & Event Index	Big Data Repository has "Event" records with inferred relationships (e.g., events linked by common "Person" or "Location" entities).	1. User performs a Global Search for "Operation Thunder." 2. In search results, double-clicks the "Events" index. 3. Selects 5 relevant event records (via right-click context menu). 4. Clicks "Add to Link-Analysis" (via right-click context menu).	A new browser tab titled "Link-Analysis" opens, displaying the selected events and their detected relationships as an interactive graph.
TC-GRPH-002	Automatic Graph Plotting	Selected data points have predefined relationships (e.g., Person A involved in	1. Initiate graph display (e.g., via TC-GRPH-001).	The system automatically renders a graph showing nodes (events,

		Event X, Event X occurred at Location Y).		persons, locations) and edges (relationships) as defined in the data.
<b>TC-GRPH-003</b>	<b>Explode Graph by Relation Node - Person</b>	Graph is displayed. Selected data points include multiple "Person" entities.	1. User selects "Explode by Relation Node" option. 2. Chooses "Person" as the criteria.	The graph re-renders or expands, highlighting and grouping nodes around common "Person" entities, clearly showing connections via individuals.
<b>TC-GRPH-004</b>	<b>Explore Connections - View Properties (RHS Panel)</b>	Graph is displayed.	1. User clicks on a specific node (e.g., a "Location" node).	A RHS (Right-Hand Side) panel opens/updates, displaying detailed properties of the selected "Location" node (e.g., coordinates, associated events, classification).
<b>TC-GRPH-005</b>	<b>Explore Connections - Expand Node</b>	Graph is displayed. A node has un-displayed connections (e.g., a "Person" node with 5 additional un-plotted event connections).	1. User clicks on a node and selects "Expand" or double-clicks it.	The graph expands to reveal the previously hidden connected nodes and their relationships, extending the visual network.
<b>TC-GRPH-006</b>	<b>Interact with Graph Elements - Drag Node &amp; Zoom</b>	Graph is displayed.	1. User clicks and drags a node to a new position. 2. User uses mouse scroll to zoom in/out on the graph.	1. The selected node moves to the dragged position without affecting its connections. 2. The graph view scales in/out, allowing detailed inspection or broader overview.
<b>TC-GRPH-007</b>	<b>Save Graph View (Export Image)</b>	A specific graph view is displayed, potentially with applied filters or layouts.	1. User selects "Export Graph" option. 2. Chooses an image format (e.g., PNG, JPEG) and clicks "Save."	An image file of the current graph visualization is downloaded or saved to the user's system.
<b>TC-GRPH-008</b>	<b>No Relationships Found</b>	Selected data points exist but have no defined relationships between them in the	1. User selects unrelated data points (e.g., an Excel record and a scanned	The graph displays individual nodes for each selected data point, but no

		Big Data Repository.	document) and adds them to "Link-Analysis."	connecting edges are drawn, and a message indicates "No relationships found."
--	--	----------------------	---	---

## Test Cases for Personalize Dashboards with Widgets

Test Case ID	Description	Preconditions	Steps	Expected Result
TC-DASH-001	Add a New Widget to Dashboard	User is logged in and on their personal dashboard.	<ol style="list-style-type: none"> <li>1. User clicks "Add Widget" button.</li> <li>2. User selects "Summary Statistics" widget from the catalog.</li> <li>3. User configures the widget (e.g., selects "Total Incidents" metric).</li> <li>4. User clicks "Add to Dashboard."</li> </ol>	The "Summary Statistics" widget appears on the dashboard, displaying the "Total Incidents" metric.
TC-DASH-002	Resize a Widget	A "Bar Chart" widget is present on the dashboard.	<ol style="list-style-type: none"> <li>1. User clicks and drags the corner/edge of the "Bar Chart" widget.</li> <li>2. Resizes it to be larger.</li> </ol>	The "Bar Chart" widget resizes correctly, and its content scales appropriately within the new dimensions.
TC-DASH-003	Rearrange Widgets via Drag-and-Drop	Two widgets, "KPI: Active Alerts" and "Recent Activity Feed," are on the dashboard.	<ol style="list-style-type: none"> <li>1. User clicks and drags "KPI: Active Alerts" widget.</li> <li>2. Moves it to a new position to the right of "Recent Activity Feed."</li> </ol>	The "KPI: Active Alerts" widget moves to the new position, and other widgets adjust their layout accordingly without overlap.
TC-DASH-004	Configure a Data-Driven Widget (Chart Type)	An "Events Trend" widget (chart) is on the dashboard.	<ol style="list-style-type: none"> <li>1. User clicks on the "Events Trend" widget's configuration icon.</li> <li>2. Changes chart type from "Line Chart" to "Area Chart."</li> <li>3. Clicks</li> </ol>	The "Events Trend" widget updates its visualization

			"Apply" or "Save."	from a line chart to an area chart.
<b>TC-DASH-005</b>	<b>Configure an Informational Widget (Content Source)</b>	A "News Feed" widget is available.	1. User adds "News Feed" widget. 2. Clicks on its configuration icon. 3. Enters a specific RSS feed URL (e.g., <a href="https://example.com/news.xml">https://example.com/news.xml</a> ). 4. Clicks "Apply" or "Save."	The "News Feed" widget displays content from the specified RSS feed.
<b>TC-DASH-006</b>	<b>Remove a Widget</b>	A "To-Do List" widget is present on the dashboard.	1. User clicks on the "To-Do List" widget's remove/delete icon. 2. Confirms removal if prompted.	The "To-Do List" widget is removed from the dashboard, and remaining widgets may adjust.
<b>TC-DASH-007</b>	<b>Automatic Saving of Dashboard Layout</b>	User adds, resizes, and rearranges several widgets without explicitly clicking a "Save" button.	1. Perform TC-DASH-001, TC-DASH-002, TC-DASH-003 actions. 2. Log out and log back in (or refresh the page).	The dashboard retains all the changes (added widgets, resized, rearranged positions) upon reload, indicating automatic saving.
<b>TC-DASH-008</b>	<b>Dashboard Responsiveness to Screen Size</b>	User views the customized dashboard on different screen sizes (e.g., desktop, tablet, resized browser window).	1. Open the dashboard on a desktop. 2. Resize the browser window to simulate a smaller screen. 3. (Optional) Open on a tablet device.	Widgets automatically adjust their size and/or layout to fit the new screen dimensions without overlapping or significant loss of readability.

## Test Cases for Generate and Manage Output Reports

Test Case ID	Description	Preconditions	Steps	Expected Result
TC-RPT-001	Generate On-Demand Report - PDF Format	Data Analyst has permission; data for "Daily Activity Report" is available.	1. Data Analyst navigates to "Report Generation" module. 2. Selects "Daily Activity Report" type. 3. Sets date range "yesterday." 4. Selects "PDF" as output format. 5. Clicks "Generate Report."	A PDF report named "Daily Activity Report [date]" is successfully generated and stored in the report repository. User is notified of successful generation.
TC-RPT-002	Generate On-Demand Report - Microsoft Word (.docx)	Data Analyst has permission; data for "Unit Strength Report" is available.	1. Data Analyst navigates to "Report Generation" module. 2. Selects "Unit Strength Report" type. 3. Selects "Microsoft Word" as output format. 4. Clicks "Generate Report."	A DOCX report named "Unit Strength Report [date]" is successfully generated and stored.
TC-RPT-003	Generate On-Demand Report - Microsoft Excel (.xlsx)	Data Analyst has permission; data for "Financial Transactions Export" is available.	1. Data Analyst navigates to "Report Generation" module. 2. Selects "Financial Transactions Export" type. 3. Selects "Microsoft Excel" as output format. 4. Clicks "Generate Report."	An XLSX report named "Financial Transactions Export [date]" is successfully generated and stored.
TC-RPT-004	Generate On-Demand Report - XML Format	Data Analyst has permission; data for "System Log Export" is available.	1. Data Analyst navigates to "Report Generation" module. 2. Selects "System Log Export" type. 3. Selects "XML" as output format. 4. Clicks "Generate Report."	An XML report named "System Log Export [date]" is successfully generated and stored.
TC-RPT-005	Generate On-Demand Report - HTML Format	Data Analyst has permission; data for "Web Dashboard Summary" is available.	1. Data Analyst navigates to "Report Generation" module. 2. Selects "Web Dashboard Summary" type. 3. Selects "HTML" as output format. 4. Clicks "Generate Report."	An HTML report named "Web Dashboard Summary [date]" is successfully generated and stored.
TC-RPT-	Workflow-Driven Report	A workflow (FR-WFA-001.5) is configured to	1. Clerk/Data Analyst monitors workflow	The "Incident Summary Report"

006	Generation	generate an "Incident Summary Report" (PDF) upon workflow completion. Workflow just completed successfully.	dashboard. 2. Verify "Incident Summary Report" is present in the report manager.	(PDF) is automatically generated and listed in the central report repository.
TC-RPT-007	Manage Reports - View List	Multiple reports (PDF, DOCX, XLSX) have been generated.	1. Data Analyst navigates to "Report Manager" interface.	A list of all generated reports is displayed, showing report name, format, generation date, and generator.
TC-RPT-008	Manage Reports - Download Report	A PDF report "Daily Activity Report [date]" exists in the Report Manager.	1. Data Analyst selects "Daily Activity Report [date]" from the list. 2. Clicks "Download."	The "Daily Activity Report [date].pdf" file is successfully downloaded to the user's local system.
TC-RPT-009	Report Generation Failure - Data Errors	Data for a report contains invalid values (e.g., non-numeric in a numeric field) which break the template.	1. Data Analyst attempts to generate a report using the problematic data.	System displays an error message (e.g., "Report generation failed due to data errors"). An error log provides details of the issue.
TC-RPT-010	Report Generation Failure - Template Issues	A report template is corrupted or incorrectly configured.	1. Data Analyst attempts to generate a report using the corrupted template.	System displays an error message (e.g., "Report generation failed: invalid template").
TC-RPT-011	Report Content Adherence to Criteria	Data Analyst generates a report with filter "Events only for Region A."	1. Data Analyst generates the report. 2. Downloads and opens the report. 3. Verifies report content.	The report content accurately reflects only events for "Region A," and no data from other regions is present.

## Test Cases for Leverage LLM for Enhanced Search & Summarization

### Preconditions:

- Data is ingested and available in the Big Data Repository (FR-BDR-001).
- The LLM is integrated, trained/fine-tuned, and operational.
- LLM has access to the indexed data (FR-SRCH-001.1).

Test Case ID	Description	Preconditions	Steps	Expected Result
--------------	-------------	---------------	-------	-----------------



TC-LLM-001	Enhanced Natural Language Search - Simple Question	Repository contains documents about "Company A's" financial performance from different sources.	1. User navigates to search interface. 2. Types: "What was Company A's revenue last quarter?" 3. Clicks search.	LLM synthesizes a concise answer about Company A's revenue, citing relevant documents (e.g., "According to Q4 report, revenue was X").
TC-LLM-002	Enhanced Natural Language Search - Complex Question with Filters	Repository contains financial audits for "Company X" over several quarters.	1. User navigates to search interface. 2. Types: "What are the key findings from the recent financial audits related to company X in the last quarter?" 3. Clicks search.	LLM provides a summary of key findings from recent audits of Company X, explicitly mentioning relevant documents and their content for the last quarter.
TC-LLM-003	Enhanced Natural Language Search - Entity-focused Question	Repository contains documents mentioning various locations and their activities.	1. User navigates to search interface. 2. Types: "Tell me about incidents reported in New Delhi." 3. Clicks search.	LLM provides a summary of incidents, events, or reports related to "New Delhi," referencing specific documents.
TC-LLM-004	Summarize Single Document	A PDF document "Project Alpha Review.pdf" (5 pages long) is in the repository.	1. User selects "Project Alpha Review.pdf". 2. Initiates "Summarize" action.	A concise, accurate summary of the 5-page document is generated and displayed. Option to save/export summary.
TC-LLM-005	Summarize Multiple Events from Timeline	5 "Events" (from FR-TIM-001) are selected from the timeline view.	1. User selects 5 event markers on the timeline. 2. Initiates "Summarize" action.	A combined summary of the 5 selected events is generated, highlighting common themes or key details.
TC-LLM-006	Summarize a Profile (e.g., Organization)	A "Company Profile" (aggregated data from various sources) is selected.	1. User navigates to "Company X" profile page. 2. Initiates "Summarize" action for the profile.	A concise summary of "Company X" is generated, drawing from all associated ingested data.
TC-LLM-007	Insufficient Information for Natural Language Search	User asks a question about a completely non-existent entity or topic (e.g., "What are the findings on Project	1. User types "What are the findings on Project ZZZ from 2030?" 2. Clicks search.	LLM responds with a message indicating it could not find sufficient information (e.g., "I

		ZZZ from 2030?").		couldn't find any relevant data on Project ZZZ for that period.").
<b>TC-LLM-008</b>	<b>Insufficient Information for Summarization</b>	User attempts to summarize a document that contains minimal or very generic content.	1. User selects a document with very little meaningful text (e.g., a blank page with only a title). 2. Initiates "Summarize" action.	LLM responds with a message indicating it could not generate a meaningful summary due to insufficient content, or provides a very short, generic summary.
<b>TC-LLM-009</b>	<b>LLM Hallucination Disclaimer</b>	After an LLM-generated summary/answer is displayed.	1. User reviews the LLM's output.	A clear disclaimer is visible near the LLM's output (e.g., "AI-generated content. Please verify critical information with original sources.").

## Test Cases for Define, Configure, and Generate Predictions

### Preconditions:

- The user has appropriate permissions to access and configure the Predictive Module.
- Historical data for the past three years (relevant to the chosen use case) is available in the Big Data Repository (FR-BDR-001).
- The system's mathematical modelling and pattern recognition components are operational.

Test Case ID	Description	Preconditions	Steps	Expected Result
<b>TC-PRD-001</b>	<b>Define New Use Case - "Infiltration Hotspot Prediction"</b>	Data Analyst has permission; System is operational.	1. Data Analyst navigates to "Predictive Analysis" module. 2. Initiates "Create New Use Case." 3. Enters "Infiltration Hotspot Prediction" as name. 4. Defines scope: "Predict likely locations of illegal border crossings." 5. Specifies "Infiltration Event" as the specific event. 6. Defines application context	"Infiltration Hotspot Prediction" Use Case is created and listed. System confirms technical feasibility.

			across "Northern Border Formation." 7. Saves Use Case definition.	
<b>TC-PRD-002</b>	<b>Configure Model - Data Selection (3 Years)</b>	"Infiltration Hotspot Prediction" Use Case defined. Historical "Border Crossing Records" are available for past 5 years in BDR.	1. Data Analyst selects "Infiltration Hotspot Prediction" Use Case. 2. Navigates to "Data Selection." 3. Selects "Border Crossing Records" data source. 4. Ensures the system automatically filters data for the last three years (or user explicitly selects this range).	The system confirms selected data (Border Crossing Records) covers the required 3-year historical period.
<b>TC-PRD-003</b>	<b>Configure Model - Pattern Recognition Parameters</b>	"Infiltration Hotspot Prediction" Use Case selected.	1. Data Analyst navigates to "Pattern Recognition Configuration." 2. Adjusts "Temporal Window" for pattern detection (e.g., "weekly patterns"). 3. Adjusts "Anomaly Threshold" (e.g., "3 standard deviations"). 4. Saves configuration.	The pattern recognition component is configured with the specified parameters for the Use Case.
<b>TC-PRD-004</b>	<b>Configure Model - Trend Identification Parameters</b>	"Infiltration Hotspot Prediction" Use Case selected.	1. Data Analyst navigates to "Trend Identification Configuration." 2. Selects "Seasonality: Monthly" and "Growth Rate: Linear." 3. Saves configuration.	The trend identification component is configured for monthly seasonality and linear growth analysis for the Use Case.
<b>TC-PRD-005</b>	<b>Configure Model - Prediction Scenario Mapping</b>	"Infiltration Hotspot Prediction" Use Case selected.	1. Data Analyst navigates to "Prediction Scenario Mapping." 2. Links identified "Increased Border Activity" patterns to "Likely Hotspots for Infiltration" prediction type. 3. Saves mapping.	The model is configured to output "Likely Hotspots for Infiltration" when "Increased Border Activity" patterns are detected.
<b>TC-PRD-006</b>	<b>Train/Build Prediction Model</b>	"Infiltration Hotspot Prediction" Use Case fully configured with data selected.	1. Data Analyst clicks "Train Model" (or system auto-trains upon configuration save). 2. Monitors training progress.	The system successfully processes historical data, builds/optimizes the predictive model, and indicates "Model Training Complete."
<b>TC-PRD-007</b>	<b>Generate Predictions - "Infiltration Hotspot"</b>	"Infiltration Hotspot Prediction" model is trained (TC-	1. Data Analyst clicks "Generate Predictions" for "Infiltration Hotspot Prediction" Use Case. 2.	Predictions for "Likely hotspots for infiltration" are generated. These

	<b>Prediction"</b>	PRD-006). Current data is available.	Monitors prediction generation status.	predictions are visible within the system (e.g., on the GIS dashboard).
<b>TC-PRD-008</b>	<b>Generate Predictions - "Unit Movement Forecast"</b>	A separate Use Case "Unit Movement Forecast" is defined and trained. Current unit status data is available.	1. Data Analyst triggers prediction for "Unit Movement Forecast."	Predictions for "Units which may go for training or move" are generated and available (e.g., on a timeline view or dedicated report).
<b>TC-PRD-009</b>	<b>Insufficient Historical Data Warning</b>	Data Analyst selects a Use Case but chooses a data source with only 1 year of historical data.	1. Data Analyst navigates to "Data Selection" for a Use Case. 2. Selects data source events_last_year.csv. 3. Attempts to proceed or train model.	System displays a warning message (e.g., "Insufficient historical data. Minimum 3 years required.") and prevents model training/prediction.

## Test Cases for Ingest IMINT Data into Form-Based Collation System

### Preconditions:

- The IMINT Analyst has appropriate permissions to ingest data.
- The system's form-based collation system is operational.
- For trial purposes, the IMINT data in its existing format is accessible.
- For operational ingestion, IMINT from the IFC workstation of IIT is accessible.

Test Case ID	Description	Preconditions	Steps	Expected Result
<b>TC-IMT-001</b>	<b>Successful Trial Ingestion - Standard Image Format</b>	IMINT Analyst has permissions. trial_image.jpg (with embedded metadata like Lat/Long, date) is accessible.	1. IMINT Analyst logs in, navigates to "IMINT Ingestion Dashboard." 2. Selects "Trial Ingestion (Existing Format)." 3. Specifies path to trial_image.jpg. 4. Clicks "Initiate Ingestion." 5. Reviews populated form. 6. Clicks "Save"	trial_image.jpg is ingested. Relevant metadata (Lat/Long, date) is extracted and correctly populated into the form-based collation system. Ingestion status is "Completed."

			Collation."	
TC- IMT- 002	<b>Successful Trial Ingestion - Non-Standard but Supported Format</b>	IMINT Analyst has permissions. trial_data.dat (a custom format, but with pre-configured flexible mapping in system) is accessible.	1. IMINT Analyst logs in, navigates to "IMINT Ingestion Dashboard." 2. Selects "Trial Ingestion (Existing Format)." 3. Specifies path to trial_data.dat. 4. Clicks "Initiate Ingestion." 5. Reviews populated form. 6. Clicks "Save Collation."	trial_data.dat is ingested. Data is extracted based on flexible mapping and correctly populated into the form-based collation system. Ingestion status is "Completed."
TC- IMT- 003	<b>Successful IFC Workstation Ingestion (Standardized)</b>	IMINT Analyst has permissions. IFC workstation is accessible and provides IMINT in the user-standardized format.	1. IMINT Analyst logs in, navigates to "IMINT Ingestion Dashboard." 2. Selects "IFC Workstation Ingestion (Standardized)." 3. (System automatically connects/prompts for details) 4. Clicks "Initiate Ingestion." 5. Reviews populated form. 6. Clicks "Save Collation."	IMINT from IFC workstation is ingested. Data is extracted according to the pre-defined standard and correctly populated into the form-based collation system. Ingestion status is "Completed."
TC- IMT- 004	<b>Ingestion Failure - Unsupported Trial Format</b>	IMINT Analyst has permissions. unsupported.xyz (a completely unknown/unreadable file format) is accessible.	1. IMINT Analyst logs in, navigates to "IMINT Ingestion Dashboard." 2. Selects "Trial Ingestion (Existing Format)." 3. Specifies path to unsupported.xyz. 4. Clicks "Initiate Ingestion."	System displays an error message (e.g., "Unsupported file format. Cannot ingest.") Ingestion status is "Failed." No data is collated.
TC- IMT- 005	<b>Ingestion Failure - Connectivity Issue to IFC Workstation</b>	IMINT Analyst has permissions. IFC workstation is offline or inaccessible (e.g., wrong IP, firewall).	1. IMINT Analyst logs in, navigates to "IMINT Ingestion Dashboard." 2. Selects "IFC Workstation Ingestion (Standardized)." 3. Clicks "Initiate	System displays a connection error message (e.g., "Failed to connect to IFC workstation. Please check network and credentials.") Ingestion status is "Failed."

			Ingestion."	
TC- IMT- 006	<b>Ingestion Failure - Standardization Mismatch from IFC</b>	IMINT Analyst has permissions. IFC workstation provides IMINT, but a critical field (e.g., "Target Type") is missing or malformed in the standardized output.	1. IMINT Analyst logs in, navigates to "IMINT Ingestion Dashboard." 2. Selects "IFC Workstation Ingestion (Standardized)." 3. Clicks "Initiate Ingestion."	System flags parsing errors for the malformed/missing data elements. The form-based collation system shows warnings or blank fields for affected data. Ingestion status may be "Completed with Errors."
TC- IMT- 007	<b>Ingestion - Insufficient Data for Core Collation Fields</b>	IMINT Analyst has permissions. partial_data.jpg is ingested, but lacks a mandatory field like "Observation Date" or "Target Name."	1. IMINT Analyst logs in, navigates to "IMINT Ingestion Dashboard." 2. Selects "Trial Ingestion (Existing Format)." 3. Specifies path to partial_data.jpg. 4. Clicks "Initiate Ingestion." 5. Reviews populated form.	System alerts the user that mandatory fields are missing or populated with "N/A." The form might be saved, but with a warning. Ingestion status is "Completed with Warnings."
TC- IMT- 008	<b>Manual Adjustment and Save of Collated Record</b>	After a successful ingestion (e.g., TC-IMT-001), the form has some minor errors or needs additions.	1. IMINT Analyst reviews the populated form (e.g., corrects a typo in "Description," adds a "Comments" field). 2. Clicks "Save Collation."	The manually adjusted data is saved, overwriting/updating the automatically extracted data in the form-based collation system.
TC- IMT- 009	<b>Monitor Ingestion Status on Dashboard - Success</b>	After a successful ingestion (e.g., TC-IMT-001).	1. IMINT Analyst views the "IMINT Ingestion Dashboard."	The dashboard clearly shows the recent ingestion of trial_image.jpg with a status of "Completed."
TC- IMT- 010	<b>Monitor Ingestion Status on Dashboard - Failure</b>	After an ingestion failure (e.g., TC-IMT-004).	1. IMINT Analyst views the "IMINT Ingestion Dashboard."	The dashboard clearly shows the attempted ingestion of unsupported.xyz with a status of "Failed" and an associated error message/link to logs.