



REQUIREMENTS CAPTURE DOCUMENT

[Product Name]

Version 0.1

Date: May 27, 2025

Prepared By:
[XYZ Company, Department]

1. Introduction

1.1 Purpose

The purpose of this document is to formally capture, define, and consolidate all functional and non-functional requirements for the Intelligence Fusion Center (IFC), Release V1.0.0. This document serves as the primary source of truth for all subsequent design, development, and testing activities, ensuring a common understanding among all stakeholders. It aligns with the "Software Requirements Analysis Process" as outlined in ISO/IEC 12207, providing a foundational input for the software lifecycle.

1.2 Scope

This document specifies the requirements for the initial release (V1.0.0) of the Intelligence Fusion Center (IFC) System. It covers functionalities related to data ingestion, processing, analysis, visualization, reporting, and predictive capabilities from diverse intelligence data sources. It does not include detailed design specifications, project management plans, or specific testing procedures, which are addressed in separate documentation.

1.3 Definitions and Acronyms

- **BDR:** Big Data Repository
- **Clerk:** User role responsible for initial data ingestion and workflow forwarding.
- **Admin:** User role responsible for system configuration, pipeline setup, classification, and dictionary management.
- **Analyst:** User role responsible for advanced search, analysis, report generation, and predictive model configuration.
- **FR:** Functional Requirement
- **IDD:** Interface Design Document
- **IFC:** Intelligence Fusion Center
- **IMINT:** Imagery Intelligence
- **ISO/IEC 12207:** International standard for software life cycle processes.
- **KPI:** Key Performance Indicator
- **LLM:** Large Language Model
- **NFR:** Non-Functional Requirement
- **OCR:** Optical Character Recognition

- **QA:** Quality Assurance
- **RCD:** Requirements Capture Document
- **RHS:** Right-Hand Side
- **RTM:** Requirements Traceability Matrix
- **SUT:** Software Under Test
- **SSCT:** System/Software/Solution/System Component Testing (as per FR-PRD-001.8)
- **UI:** User Interface
- **UX:** User Experience
- **V&V:** Verification and Validation

1.4 References

- RFP - Intelligence Fusion Center (IFC) System
- ISO/IEC 12207: Systems and software engineering — Software life cycle processes
- ISO/IEC/IEEE 29119: Software and systems engineering — Software testing

2. Overall Description

2.1 Product Perspective

The Intelligence Fusion Center (IFC) System is envisioned as a standalone, integrated intelligence platform within a broader organizational intelligence ecosystem. It acts as a central hub for consolidating, processing, and analyzing diverse data sources, reducing manual effort and enhancing decision-making. It integrates with existing organizational data sources (e.g., internal databases, file shares) and provides user interfaces for various roles, but it is not dependent on other *new* systems for its core functionality.

2.2 Product Functions (High-Level Capabilities)

The IFC System provides the following high-level capabilities:

- **Data Ingestion:** Automated and on-demand ingestion from various structured and unstructured data sources.
- **Intelligent Processing:** OCR for scanned documents, and text analytics for classification and entity extraction.
- **Unified Data Repository:** Centralized storage and automatic co-relation of all ingested data.

- **Advanced Search:** Efficient retrieval of information using various search paradigms, including natural language queries.
- **Workflow Automation:** Definition and execution of multi-stage data processing workflows with human review and approval.
- **Multi-Dimensional Analysis:** Geospatial, Temporal, and Graph-based visualization and analysis.
- **Customizable Dashboards:** Personalized display of key performance indicators and summaries via widgets.
- **Comprehensive Reporting:** Generation of reports in multiple formats, including adherence to standardized templates.
- **Predictive Analytics:** Generation of forecasts and identification of trends for specific use cases.
- **IMINT Integration:** Specialized ingestion and collation of Imagery Intelligence data.
- **System Configuration & Management:** Tools for managing data classifications, abbreviations, and ingestion pipelines.

2.3 User Characteristics

The IFC System will be used by different user roles, each with specific needs and technical proficiencies:

- **Clerk:**
 - **Proficiency:** Basic computer literacy, familiarity with file systems and data entry.
 - **Tasks:** Initiating data ingestion, forwarding workflow tasks, basic monitoring.
 - **Needs:** Intuitive interfaces, clear status indicators, minimal configuration.
- **Analyst:**
 - **Proficiency:** Moderate to high technical proficiency, understanding of data structures, analytical thinking.
 - **Tasks:** Performing advanced searches, conducting multi-dimensional analysis (GIS, Timeline, Graph), generating reports, configuring predictive models, leveraging LLM for insights.

- **Needs:** Powerful query tools, interactive visualizations, customizable reporting, access to raw and processed data.
- **Admin:**
 - **Proficiency:** High technical proficiency, understanding of database concepts, system configuration.
 - **Tasks:** Configuring data ingestion pipelines, managing data classifications, augmenting dictionaries, setting up workflow rules, managing report access.
 - **Needs:** Robust configuration interfaces, comprehensive logging, error management tools, access control.
- **IMINT Analyst:**
 - **Proficiency:** Specialized knowledge in Imagery Intelligence, understanding of IMINT formats.
 - **Tasks:** Ingesting IMINT data, reviewing collated IMINT records.
 - **Needs:** Dedicated IMINT ingestion dashboard, support for specific IMINT formats, form-based collation.
- **Workflow Approver:**
 - **Proficiency:** Moderate computer literacy, strong domain knowledge for decision-making.
 - **Tasks:** Reviewing intermediate and final outputs of workflows, approving or rejecting tasks.
 - **Needs:** Clear presentation of data for review, audit trails of changes, easy navigation to reference files.

2.4 General Constraints

- **Operational Environment:**
 - Deployment Environment: [e.g., On-premises servers]
 - Operating System: [e.g., Linux-based server, Windows client]
 - Network: [e.g., Intranet/VPN access, specific bandwidth requirements]
- **Regulatory & Compliance:**
 - Data Classification: Adherence to organizational data classification policies (e.g., Confidential, Restricted).

- Audit Trails: All user actions and system changes must be logged for auditability.
- Data Retention: Compliance with data retention policies.
- **Technical Constraints:**
 - Big Data Handling: Ability to process and store terabytes of data.
 - LLM Integration: Dependence on internal LLM Module/APIs.
 - OCR Accuracy: Specified accuracy targets for multilingual OCR.
 - Predictive Model Accuracy: Minimum 30% accuracy at SSCT.
- **User Interface:**
 - Web-based interface compatible with modern browsers.
 - Responsive design for various device types.
 - Adherence to internal UI/UX guidelines (if applicable).
- **Performance:** All critical user interactions and system processes must meet defined performance targets (see NFRs).

2.5 Assumptions and Dependencies

- **Data Source Accessibility:** Assumed that external data sources (Excel, Access, SQL DBs, IFC workstation) are accessible to the system via network and appropriate credentials.
- **Existing IMINT Formats:** Assumed that user-provided IMINT formats (for trial) are well-defined and consistently structured or can be flexibly mapped.
- **LLM API Availability:** Assumed that the internal LLM API is consistently available, performs reliably, and meets response time expectations.
- **Third-Party Software/Libraries:** Assumed that necessary third-party libraries (e.g., for PDF/DOCX parsing, GIS mapping, database drivers) are available and compatible.
- **Network Infrastructure:** Assumed stable and sufficient network bandwidth for data transfers and application access.
- **Security Policies:** Assumed that network security policies allow necessary connections for data ingestion and external API calls.

3. Specific Functional Requirements

This section details the explicit functional capabilities the IFC System must provide. Each requirement is unique, verifiable, and traceable.

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 document images.
- **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 multiple languages including English, Hindi, Mandarin etc.
- **FR-OCR-001.3:** 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 (e.g., linking event ID across Excel and SQL data).

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 (e.g., document, events, allorbat).
- **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 as a tree-map and a list.

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, cleansing, 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 (Location, Orbat, Abbreviation data for augmentation).
- **FR-TA-001.4:** The text analytics module shall perform entity extraction to identify and categorize key entities (e.g., classifications and subtypes, units, locations, events, 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 (e.g., vector layer, raster layer).
- **FR-GIS-001.3:** Users shall be able to filter and interact with geospatial data directly on the map (e.g., event date, classification, subtype).
- **FR-GIS-001.4:** The system shall support plotting custom points, events, areas (polygons), and routes on the map. On click it should also show associated information of that events. These elements can be ingested, saved and exported (in KML, SHP, and CSV formats).
- **FR-GIS-001.5:** The system shall provide Geo-Fencing and Geo-Search capabilities, allowing users to mark areas of interest (polygon, circle, square) to filter/search data (events, raster, vector layers) for specific map types (cluster, pushpin, bubbles).
- **FR-GIS-001.6:** The system shall create a timeline layer on top of the geo-layer to determine temporal-geospatial relationships.
- **FR-GIS-001.7:** The system shall overlay visual representations of reports or analytical findings that have a geospatial component (e.g., heatmaps of activity, clusters, pushpins, timeline, synchronization maps).
- **FR-GIS-001.8:** 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 by date range or other attributes.
- **FR-TIM-001.4:** The timeline shall display event details upon clicking or hovering on an event marker.

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 (Link-Analysis).
- **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.
- **FR-GRPH-001.5:** The system shall allow users to export graph views for future reference or sharing.

3.1.10 Predictive Module (FR-PAN-001)

- **FR-PAN-001.1:** The system shall provide the ability to generate predictions based on analyzed trends and patterns for user-defined Use Cases.
- **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 Stage.

3.1.11 IMINT Module (FR-IMT-001)

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

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, charts (bar, pie, line).
- **FR-WGT-001.3:** Admin Users shall be able to add, remove, resize, and rearrange widgets on their personalized dashboards.

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, downloading, and archiving.

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 LLM shall be capable of summarizing user-specified events from ingested data.
- **FR-LLM-001.3:** The LLM 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 LLM.
- **FR-LLM-001.5:** The LLM 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.

4. Non-Functional Requirements (NFRs)

4.1 Security Requirements

- **NFR-SEC-001:** The system shall support role-based access control (RBAC) to ensure users can only access features and data for which they have explicit permissions.
- **NFR-SEC-002:** All sensitive data (e.g., classified documents, personal information) shall be encrypted at rest and in transit.
- **NFR-SEC-003:** The system shall provide robust authentication mechanisms for user login.
- **NFR-SEC-004:** The system shall log all access attempts and data modifications for auditing purposes.

4.3 Usability Requirements

- **NFR-USG-001:** The user interface shall be intuitive and easy to navigate for all user roles.
- **NFR-USG-002:** The system shall provide clear, concise, and context-sensitive help documentation for all functionalities.
- **NFR-USG-003:** All user input forms shall provide inline validation and informative error messages.
- **NFR-USG-004:** Dashboard widgets and analytical modules shall support drag-and-drop and point-and-click interactions where appropriate.

4.4 Reliability Requirements

- **NFR-REL-001:** The system shall be highly available.
- **NFR-REL-002:** The system shall implement robust error handling and logging mechanisms to capture and report all system failures.
- **NFR-REL-003:** In the event of a system failure, data integrity shall be maintained, and the system shall recover to a consistent state.
- **NFR-REL-004:** The system shall support backups of the Big Data Repository and configurations every week/month securely.
- **NFR-REL-005:** The system shall implement encryption of data at rest and transit.

4.5 Maintainability Requirements

- **NFR-MTN-001:** The system shall be modular in design, allowing for independent updates and maintenance of components (e.g., OCR engine, LLM integration).

- **NFR-MTN-002:** The codebase shall adhere to established coding standards and be well-documented.
- **NFR-MTN-003:** The system shall provide comprehensive logging and monitoring capabilities to assist in troubleshooting.

4.6 Portability Requirements

- **NFR-PTY-001:** The system shall be deployable on standard Linux-based server distributions (e.g., Ubuntu LTS).
- **NFR-PTY-002:** The web-based user interface shall be compatible with the latest stable versions of Chrome, Firefox, and Edge browsers.

4.7 Scalability Requirements

- **NFR-SCL-001:** The Big Data Repository shall be horizontally scalable to accommodate growth in data volume and concurrent users.
- **NFR-SCL-002:** The search index shall scale to maintain efficient retrieval times as data volume increases.

5. Interface Requirements

5.1 User Interfaces

- **REQ-UI-001:** The system shall provide a web-based Graphical User Interface (GUI) accessible via standard web browsers.
- **REQ-UI-002:** The GUI shall support responsive design principles to adapt to desktop, tablet, and mobile screen sizes.
- **REQ-UI-003:** User roles (Clerk, Data Analyst, Data Administrator, IMINT Analyst, Workflow Approver) shall have distinct dashboards and access privileges.
- **REQ-UI-004:** The GUI shall provide visual feedback for all user actions (e.g., loading indicators, success messages, error alerts).

5.2 Software Interfaces

- **REQ-SI-001:** The system shall integrate with external SQL databases (e.g., PostgreSQL, MySQL, SQL Server) for data ingestion.
- **REQ-SI-002:** The system shall utilize an external search engine (e.g., Elasticsearch) for indexing and retrieval.
- **REQ-SI-003:** The system shall integrate with a Large Language Model (LLM) API (e.g., Google Gemini API) for natural language processing tasks (summarization, natural language search).

- **REQ-SI-004:** The system shall integrate with an OCR engine (e.g., Tesseract or commercial SDK) for text extraction from images.
- **REQ-SI-005:** The system shall utilize GIS mapping libraries and APIs (e.g., Google Maps API, Leaflet/OpenLayers, PostGIS) for geospatial visualizations.
- **REQ-SI-006:** The system shall support programmatic interfaces (APIs) for importing and exporting data in formats like KML, SHP, and CSV for custom map elements.
- **REQ-SI-007:** The system shall interface with Microsoft Office file formats (.docx, .xlsx, .accdb) for ingestion and report generation.

5.3 Communications Interfaces

- **REQ-CI-001:** The system shall support standard network protocols (e.g., HTTP/S, TCP/IP, SMB/NFS) for data source connectivity.
- **REQ-CI-002:** The system shall securely communicate with external APIs (LLM, GIS) using HTTPS and API key authentication.
- **REQ-CI-003:** The system shall provide secure network access to designated IFC Workstation(s) for IMINT data ingestion.

6. Data Requirements

6.1 Data Model

- **REQ-DM-001:** The Big Data Repository shall support a unified logical data model capable of representing structured, semi-structured, and unstructured data from various sources.
- **REQ-DM-002:** The data model shall support relationships between entities across different data sources (e.g., Event ID, Person, Location, Organization).
- **REQ-DM-003:** All ingested data shall be tagged with its original source metadata (e.g., filename, ingestion date, source type).
- **REQ-DM-004:** Extracted classifications and entities from text analytics shall be stored as structured metadata associated with the original documents.

6.2 Data Integrity

- **REQ-DI-001:** The system shall maintain data integrity during ingestion, processing, and storage, ensuring no data loss or corruption.
- **REQ-DI-002:** The system shall provide mechanisms for conflict resolution during data collation (e.g., "most recent wins," manual review).

- **REQ-DI-003:** Data validation rules shall be applied during ingestion to ensure data quality.

6.3 Data Retention

- **REQ-DR-001:** The system shall support configurable data retention policies based on classification or data type.
- **REQ-DR-002:** Archived data shall remain retrievable as per defined policies.

7. Traceability

All requirements defined in this document will be traced throughout the system's lifecycle using a Requirements Traceability Matrix (RTM). The RTM will map each requirement to relevant design elements, code modules, and test cases, ensuring comprehensive coverage and verification. The RTM (RTM-IFC-V1.0) is a separate deliverable that ensures this traceability.

8. Approvals

This Requirements Capture Document has been reviewed and approved by the undersigned, signifying their agreement with the defined requirements for the Intelligence Fusion Center (IFC), Release V1.0.0.

Role	Name	Signature	Date
Project Manager	[Manager's Name]		May 25, 2025
Lead System Engineer	[Lead System Engineer's Name]		May 25, 2025
Product Owner	[PO's Name]		May 25, 2025
QA Lead	[QA Lead's Name]		May 25, 2025
