Client Requirements Document

Court Case Analytics & Case Flow Intelligence Platform

---

1. Business Background

The client operates across multiple courts and jurisdictions, each maintaining independent case management systems. Due to the lack of a centralized data platform, the client faces challenges in obtaining timely and consistent insights into case volumes, case lifecycle timelines, and court-level performance metrics.

Existing reporting processes are largely manual, inconsistent across jurisdictions, and delayed, limiting the ability of leadership to proactively identify bottlenecks, manage case backlogs, and optimize judicial resource allocation.

The client requires a centralized analytics platform to consolidate court data, standardize metrics, and support data-driven decision-making.

---

2. Business Objectives

The objectives of this initiative are to:

- Centralize court case data from multiple jurisdictions into a single analytics platform.
- Provide near real-time visibility into case volumes and trends.
- Track case lifecycle metrics from filing through disposition.
- Identify bottlenecks by court, case type, and jurisdiction.
- Enable judge-level and court-level workload analysis.
- Support long-term planning and backlog reduction initiatives.
- Ensure scalability, auditability, and analytics readiness.

---

3. Functional Requirements

3.1 Data Sources

The platform must ingest data from the following sources:

1. Court case management systems:
    - Case filings
    - Case status updates
    - Hearing schedules
    - Case dispositions
2. Reference and metadata sources:
    - Court and jurisdiction mappings
    - Judge identifiers
    - Case types and categories
    - Filing and disposition codes

3.2 Data Ingestion & Processing

1. Support both batch and incremental ingestion patterns.
2. Implement a Medallion architecture:
    - **Bronze Layer:** Raw ingested data with minimal transformation
    - **Silver Layer:** Cleaned, standardized, and validated datasets
    - **Gold Layer:** Curated, analytics-ready datasets
3. Handle:
    - Schema evolution as new case attributes are introduced
    - Late-arriving or corrected case updates

    - Duplicate and inconsistent case records

3.3 Data Modeling

1. Design an analytics-optimized star schema.
2. Core entities include:
    - Fact tables capturing case lifecycle events and durations
    - Dimension tables for courts, judges, case types, and time
3. Support:
    - Historical tracking of case status changes
    - Historical tracking of judge assignments and court mappings
    - Time-based trend analysis and comparisons across courts

---

3.4 Analytics & Reporting

The platform must support analytical use cases including:

- Active case counts by court and jurisdiction
- Case backlog trends over time
- Average case duration by case type
- Filing vs disposition rate comparisons
- Judge-level caseload and throughput analysis

Dashboards must support:

- Drill-down from jurisdiction to court to judge
- Filtering by case type, status, and time period

3.5 Orchestration & Automation

1. Automate end-to-end data pipelines including:
    - Source ingestion
    - Transformation and validation
    - Analytics layer refreshes
2. Pipelines must:
    - Be idempotent and fault-tolerant
    - Support reprocessing and backfills
    - Provide monitoring, logging, and alerting

3.6 Data Quality & Validation

The system must implement data quality checks including:

- Detection of missing or invalid dates
- Identification of duplicate case identifiers
- Validation of logical case status transitions
- Enforcement of referential integrity across dimensions

Data quality metrics should be generated as part of the Silver layer processing.

3.7 Security & Governance

1. Implement role-based access control for:
    - Administrators
    - Analysts
    - Read-only users

2. Ensure:
    - Clear data lineage from raw ingestion to reporting
    - Auditability of all transformations
    - Separation of raw, cleaned, and curated datasets

---

4. Deliverables

The deliverables for this project include:

- End-to-end data platform implementing Medallion architecture
- Analytics-ready Gold tables supporting court KPIs
- Interactive dashboards and analytical views
- Automated and monitored data pipelines
- Complete project documentation including:
    - Architecture diagrams
    - Data models
    - Pipeline workflows
    - Source-to-target mappings
    - Version-controlled code repository

---

5. Success Criteria

The project will be considered successful if:

- Leadership can access timely and reliable analytics.
- Data pipelines run automatically with minimal manual intervention.
- Historical case changes are fully traceable.
- The platform scales to additional courts and data sources.
- Schema changes do not disrupt existing analytics.