

From: Robert Chen – Director of Operations, Metro Legal Services Alliance To: Data Engineering Solutions Partner
Subject: Development of Real-Time Legal Representation Gap Analysis Platform

1. Business Background

Metro Legal Services Alliance (MLSA) oversees legal representation coordination across 5 criminal courts in the metropolitan area districts. We face critical challenges in ensuring constitutional compliance with the right to counsel, particularly during high-volume periods such as arraignment surges.

We lack a centralized, real-time data system to monitor representation gaps, prioritize urgent cases, and allocate legal aid resources effectively across demographics and geographic areas.

2. Business Objectives

- **Monitor representation gaps** to ensure constitutional compliance
- **Identify high-priority cases** requiring immediate legal aid intervention
- **Enable demographic and geographic analysis** for equitable resource allocation
- **Automate alerts** for critical cases approaching hearing dates without representation

3. Functional Requirements

1. Data Sources a. Real-time court proceeding data from NYC court management systems b. Attorney assignment data from public defender databases
c. Demographic and income data for defendants d. Court scheduling and hearing information

2. Data Processing & Storage a. Store data in **Medallion architecture** (Bronze → Silver → Gold) b. Handle **schema evolution** when new legal or demographic attributes are introduced c. Implement **Slowly Changing Dimension Type 2 (SCD2)** for defendant status and case progression tracking d. Implement **Star schema** for optimized analytical queries

3. Analytics a. Use **Azure Synapse** for complex analytical queries b. Build dashboards in **Power BI** with KPIs such as: i. Representation gap percentage by court district ii. Priority case distribution by charge severity and income level iii. Time-to-representation metrics iv. Geographic heat maps showing underserved areas

4. Orchestration & Automation a. Use **Azure Data Factory** to automate: i. Real-time streaming from court systems via Event Hub ii. Daily batch processing for demographic updates iii. Gold-layer refresh for executive dashboards

5. Data Quality a. Simulate realistic data quality issues (e.g., missing defendant information, inconsistent attorney status, future hearing dates) and handle them in Silver layer processing b. Implement data validation rules for critical legal compliance fields

6. Security & Compliance a. Role-based access control for different CJA departments and legal aid organizations b. Ensure compliance with criminal justice data handling regulations c. Audit logging for all data access and case interventions

4. Deliverables

- Fully functional Azure-based real-time data pipeline
- Power BI dashboard connected to live Synapse analytics
- Automated alerting system for urgent representation needs
- Data quality validation and monitoring reports
- Full project documentation (architecture, data models, compliance procedures)

5. Success Criteria

- Dashboards provide real-time visibility into representation gaps across all 5 NYC court districts
- Automated alerts trigger within 15 minutes for critical cases (felony charges without representation)
- System accurately identifies 95%+ of defendants requiring legal aid intervention
- Response time from case identification to attorney assignment reduces by 40%

6. Expected Business Impact

- **Constitutional Compliance:** Ensure right to counsel is protected across all demographic groups
- **Resource Optimization:** Data-driven allocation of legal aid resources based on geographic and demographic need
- **Equity Improvement:** Identify and address representation disparities by income, age, and location
- **Operational Efficiency:** Reduce manual case review time and improve intervention response rates

7. Technical Architecture Requirements

- **Real-time Processing:** Azure Event Hub for streaming court data
- **Data Lake:** Azure Data Lake Storage with medallion architecture
- **Analytics Engine:** Azure Databricks for complex transformations
- **Data Warehouse:** Azure Synapse for star schema implementation
- **Visualization:** Power BI for executive and operational dashboards
- **Orchestration:** Azure Data Factory for automated pipeline management

This platform will enable Metro Legal Services Alliance to fulfill its constitutional mandate while ensuring equitable access to legal representation across all communities and case types.