**Data Quality & Integrity Assessment**

**Title:** *Audit Report on Network Data Sources and Telemetry Accuracy*

**A. Objective:**

Evaluate accuracy, completeness, timeliness, and reliability of data streams used for monitoring and optimization in the 5G network.

**B. Data Sources Audited**

| **Source** | **Type** | **Description** | **Issues Identified** | **Severity** | **Recommended Actions** |
| --- | --- | --- | --- | --- | --- |
| Network Element Logs | Event logs from RAN, Core network, SDN devices | Missing timestamps, incomplete logs | Medium | Standardize log formats, sync NTP clocks |  |
| Performance Counters | KPIs from EMS/NMS platforms (throughput, latency, errors) | Data gaps during peak loads | High | Increase polling frequency, buffer overloads |  |
| Telemetry Streams | Real-time telemetry from probes and AI agents | Data noise, outliers due to sensor faults | Medium | Implement data cleaning and smoothing |  |
| OSS/BSS Databases | Customer provisioning, billing, SLA management | Data mismatch between systems | High | Implement data reconciliation processes |  |
| AI Model Inputs | Historical datasets used for SON and optimization | Labeling errors, incomplete datasets | Critical | Data validation pipelines & augmented datasets |  |
| User Experience Data | QoE reports from customer apps | Sampling bias towards urban users | Medium | Expand sampling footprint & stratified sampling |  |

**C. Assessment Results Summary**

| **Quality Dimension** | **Score (0–10)** | **Comments** |
| --- | --- | --- |
| Accuracy | 7 | Mostly reliable, with sporadic gaps |
| Completeness | 6 | Missing data in rural telemetry streams |
| Timeliness | 8 | Near real-time for core KPIs |
| Consistency | 7 | Some schema mismatches across platforms |
| Integrity | 8 | Good controls on data integrity |

**D. Recommendations**

* **Deploy centralized time synchronization** (NTP/PTP) across all network elements.
* **Enhance data ingestion pipelines** with automated error detection and correction.
* **Implement unified telemetry schema** using standards like OpenConfig and gNMI.
* **Regularly audit and validate AI training datasets** to prevent bias and drift.
* **Expand user feedback collection tools** in underserved and rural areas.