Here’s a sample Runbook / Playbook for Acceldata Data Observability Platform. This guide assumes Acceldata is being used to monitor data pipelines, infrastructure, and quality across modern data stacks like Spark, Hive, Airflow, Snowflake, etc.

## **Acceldata Data Observability Platform - Runbook / Playbook**

### **1. Overview**

This runbook outlines procedures for:

* Monitoring data pipelines
* Responding to alerts and incidents
* Troubleshooting common issues
* Performing routine maintenance tasks

### **2. System Components**

Acceldata typically monitors the following:

* Data Reliability: Schema changes, null checks, distribution checks
* Pipeline Monitoring: Airflow, Spark, Hive job performance
* Resource Monitoring: CPU, Memory, Disk usage for clusters
* DataLineage & Governance: Track flow and transformation of data
* Data Quality Rules: Thresholds, anomaly detection

### **3. Monitoring & Alerts**

#### 3.1 Dashboards

* Check Dashboards: Access via the Acceldata UI under "Dashboards"
* Dashboards include:
  + System health
  + Data pipeline SLAs
  + Data quality metrics
  + Resource usage trends

#### 3.2 Alerts

* Configure alerts via the Alerting section
* Channels: Email, Slack, PagerDuty, Webhook
* Alert Types:
  + SLA breaches
  + Data anomalies (null spike, outliers)
  + Job failures
  + Schema drifts

### **4. Incident Response**

#### 4.1 SLA Breach

* Check: DAG or job execution time on timeline view
* Action: Review logs of the job, verify resource usage
* Escalation: Notify engineering if consistent breach pattern is observed

#### 4.2 Data Quality Violation

* Check: Rule failure details (e.g., null %, duplicates)
* Action: Trace root cause via data lineage view
* Fix: Coordinate with source or transformation owners

#### 4.3 Schema Drift

* Check: Drift alert and impacted columns
* Action: Update downstream schema if expected, otherwise roll back
* Fix: Communicate schema change policy to data producers

### **5. Troubleshooting**

#### 5.1 Job Not Appearing in Acceldata

* Check: Integration connectors (Airflow/Spark/Hive) are enabled
* Action: Restart Acceldata agent if needed
* Logs: /var/log/acceldata/agent.log

#### 5.2 Data Not Refreshing in UI

* Check: ETL execution time and connector health
* Action: Validate last run time; re-trigger ETL sync

#### 5.3 Alert Flooding

* Check: Alert threshold settings and rules
* Action: Adjust sensitivity or silence non-critical alerts

### **6. Maintenance Tasks**

#### Daily

* Review overnight job and data quality alerts
* Validate system health dashboard

#### Weekly

* Audit alerting rules and adjust thresholds
* Review data lineage updates

#### Monthly

* Clean up stale data assets or deprecated pipelines
* Review integration health and connector upgrades

### **7. Admin Tasks**

#### User Management

* Navigate to Admin > Users
* Assign roles: Admin, Viewer, Editor

#### Rule Management

* Add/edit/delete quality rules via “Rules Engine”
* Example: NULL % < 5% for column ‘customer\_id’

#### Integration Management

* Supported tools: Snowflake, BigQuery, Hive, Spark, Airflow, etc.
* Validate connector health status regularly

### **8. Log & Audit Trails**

* All platform activity is logged
* Use the Audit Trail UI or export logs
* Useful for debugging, compliance, and change tracking

### **9. Escalation Matrix**

| Severity | Description | SLA | Escalation Contact |
| --- | --- | --- | --- |
| Critical | Pipeline Down, Data Loss | 1 hour | DataOps Team |
| High | SLA Breach, Quality Issue | 4 hours | Data Engineering |
| Medium | Minor Schema Drift | 1 day | Platform Owner |
| Low | Cosmetic or UI Bugs | 3 days | Acceldata Support |

### **10. Backup & Disaster Recovery**

* Daily backup of configuration and rules
* Periodic snapshots of monitored metadata
* Validate DR procedures quarterly

Would you like this turned into a formatted PDF or Confluence-style page? Also, let me know if your team has specific tools integrated (e.g., Snowflake, GCP, Airflow), and I can tailor this further.

4o

##### You said:

we are bank using cloud composer for data pipeline can you tailor

###### ChatGPT said:

Updated saved memory

Great! Since you're a bank using Cloud Composer for your data pipelines, here's a tailored runbook/playbook for Acceldata Data Observability Platform, optimized for financial data, compliance needs, and GCP-native workflows.

## ✅ **Acceldata Runbook – Bank (GCP + Cloud Composer Focus)**

### **1. Overview**

This runbook is designed for banking data teams using:

* Google Cloud Composer (Airflow-managed)
* BigQuery, GCS, Pub/Sub as core data services
* Acceldata for observability, data quality, and governance

### **2. Key Observability Coverage**

| Area | Observability by Acceldata |
| --- | --- |
| Data Pipelines (Airflow/Composer) | DAG runs, failures, latency, retries |
| Data Quality | Rule-based validation (nulls, duplicates, thresholds) |
| Data Freshness | Last load timestamps, SLA tracking |
| Schema Drift | BigQuery schema monitoring |
| Resource Usage | GKE cluster, Composer environment health |
| Lineage & Governance | Trace from GCS → BigQuery → Downstream consumers |

### **3. Monitoring Setup**

#### 3.1 Connect Cloud Composer

* Use Acceldata’s Airflow connector
* Connect via REST API or Metadata DB access
* Enable Composer metadata export for DAG visibility

#### 3.2 Connect BigQuery

* Configure BigQuery connector with service account credentials
* Monitor:
  + Query execution time
  + Table freshness
  + Row growth trends
  + Partition usage

#### 3.3 Enable Data Quality

* Define rules per table (critical for financial data):
  + NULL % < 0.5% for "transaction\_id"
  + amount > 0
  + Date consistency, duplicates, referential integrity

#### 3.4 Alert Routing

* Setup alerts via:
  + Email (DLs preferred: dataops-alerts@bank.com)
  + Slack (dedicated #data-alerts channel)
  + PagerDuty (for critical pipeline failures)

### **4. Daily Operational Playbook**

| Time | Task |
| --- | --- |
| 8:00 AM | Check Airflow DAG status in Acceldata |
| 9:00 AM | Review data quality alerts on financial tables |
| 11:00 AM | Validate schema consistency for critical BQ tables |
| 3:00 PM | Investigate anomalies flagged in dashboards |
| 5:00 PM | Review SLA dashboard and send end-of-day summary |

### **5. Incident Response Scenarios**

#### 🛑 **Critical DAG Failure (Cloud Composer)**

* Trigger: DAG failure alert
* Action:
  1. Check Acceldata → Airflow pipeline status
  2. View Composer logs via GCP UI
  3. If related to a GCS file/input delay, escalate to upstream source team
  4. Manually rerun the DAG (if safe)

#### ⚠️ **Data Quality Alert (e.g., NULL spike in** transactions**)**

* Trigger: Acceldata DQ rule failure
* Action:
  1. Use data lineage view to find the source
  2. Pull recent load audit logs from BigQuery
  3. Check GCS file validity or Pub/Sub message formats

#### 🔄 **Schema Drift in BigQuery Table**

* Trigger: Column added/removed
* Action:
  1. Validate against source schema change request
  2. If unauthorized, revert or patch DAGs with new schema handling

### **6. Monthly Maintenance**

* Audit Acceldata alert rules for false positives
* Review DAG SLAs and actual runtimes
* Clean up unused rules and stale assets
* Cross-check compliance coverage (e.g., PII tagging)

### **7. Compliance & Governance**

#### 7.1 Regulatory Readiness

* Acceldata lineage supports:
  + GDPR: Identify where PII data flows
  + SOX: Auditable trail of job executions and data changes
  + BCBS 239: Traceability from source to report

#### 7.2 Audit Logs

* Available via Acceldata or GCP logging
* Use for:
  + Incident review
  + Data quality exception justification
  + Operational traceability

### **8. Escalation Matrix**

| Issue Type | First Response | Escalation |
| --- | --- | --- |
| DAG Failures | DataOps On-call | Data Engineering |
| Schema Drift | Platform Owner | Data Governance |
| Resource Limits | GCP Infra Team | DevSecOps |
| Alert Noise | Acceldata Admin | Vendor Support |

### **9. Logs & Debugging Paths**

| Component | Path |
| --- | --- |
| Cloud Composer Logs | GCP Console > Logging > composer-\* |
| DAG Metadata | Acceldata > Airflow > DAGs |
| Acceldata Agent Logs | /var/log/acceldata/agent.log |
| Data Quality Failures | Acceldata > Data Reliability |

### **10. DR & Backup**

* Cloud Composer: Backed by GKE + Cloud Storage snapshots
* Acceldata: Config backup and metadata snapshots
* Validate Acceldata rule exports monthly

Would you like this packaged as a PDF, Confluence page, or both? Also happy to help draft Acceldata alert rules or DAG health check sensors if needed.