**Business Requirements Document (BRD)**

**Project:** DigitizeMI – Internal Reporting Automation  
**Prepared For:** Group Finance Internal Reporting  
**Prepared By:** Veerabadran Sudhakar  
**Date:** 09/30/2025  
**Version:** 1.0

**1. Executive Summary**

The Internal Reporting process today is heavily dependent on manual workflows, SAP BO extracts, Excel/EUDAs, and Alteryx transformations. This model introduces operational inefficiencies, inconsistent logic, high risk of errors, and compliance concerns.

The DigitizeMI program will automate the end-to-end MI reporting pipeline:

* Ingesting data strategically from authoritative sources and CFO Dial-In/ROCP uploads for adjustments and mappings information.
* Securing sensitive data through Protegrity tokenization and Immuta policy enforcement.
* Replacing EUDA (using GenAI App: TranslateX) and Alteryx workflows (using GitLab Duo) with standardized SQL and dbt transformations.
* Delivering governed, trusted data through S3 Iceberg tables consumed by Redshift, Snowflake, and Databricks.
* Providing analytics via Tableau (repointed), Power BI (onboarded), and SAC, all accessed through the CFO Dial-In UI as the single launchpad.

This BRD outlines scope, objectives, requirements, detailed Phase 1 plan, target vision, risks, testing, and governance.

2. Project Scope

In Scope

* Three priority MI reports:
  + Client Revenue Landing Page
  + Cross Sell
  + Trend Analysis
* Automated ingestion pipelines from operational data sources.
* Controlled adjustments and mappings via OCP portal.
* Tokenization and de-tokenization framework.
* Replacement of EUDA and Alteryx logic with governed SQL.
* Redshift semantic views and final tables for Phase 1 reporting.
* Tableau dashboards repointed to governed sources.
* Onboarding of Snowflake, Databricks, dbt, Airflow, Immuta, Alation, Power BI.
* CFO Dial-In UI enhancements for uploads, writeback, monitoring, and dashboards.

Out of Scope

* Non-finance reporting domains.
* SAP BO modernization beyond minimum extracts for transition.

3. Objectives & Success Criteria

Objectives

* Automation: Replace manual EUDA/Alteryx processes with governed pipelines.
* Governance: Centralize metadata, lineage, and stewardship.
* Security: Enforce tokenization and policy-driven access.
* Scalability: Enable future consumption via Snowflake and Databricks.
* User Experience: Provide CFO UI as a single launchpad for data, commentary, and reporting.

Success Criteria

* 100% elimination of EUDA/Alteryx logic for in-scope reports.
* ≥99.9% reconciliation accuracy vs legacy outputs.
* Tokenization at ingestion; no cleartext persisted at rest.
* All datasets registered in Lake Formation and visible in Alation lineage.
* Tableau reports repointed and validated; Power BI dashboards delivered.
* CFO UI adoption by ≥80% of target user base in 90 days.

4. Stakeholders

* Product Owner: Scope and business acceptance.
* Data Engineering: Ingestion, transformations, dbt models.
* Platform Engineering: Redshift, Snowflake, Databricks onboarding.
* Security & Governance: Tokenization, Immuta, Lake Formation, Alation.
* BI & Reporting: Tableau migration, Power BI build.
* CFO UI Team: Launchpad integration.
* QA & Testing: Data reconciliation, UAT, performance testing.
* PMO: Program tracking, dependencies, and reporting.

5. Dependencies

* Authoritative sources available and approved.
* TranslateX and GitLab Duo GenAI patterns accessible.
* OCP portal supports controlled templates and stewardship.
* Protegrity and Immuta integrated with EDP.
* Snowflake and Databricks tenancy provisioned.
* Lake Formation and Alation fully enterprise-enabled.

6. Non-Functional Requirements

* Performance: End-to-end pipeline loads complete by T+1 7 AM ET; report query P95 ≤ 5 seconds.
* Resilience: Retry/recovery ≤ 15 minutes for transient failures.
* Scalability: 3× data growth supported without redesign.
* Compliance: SOX/GDPR adherence; immutable audit logs.
* Security: Tokenization at write; JIT de-tokenization only.
* Usability: CFO UI provides intuitive, role-aware experience.

7. Phase 1 – Near-Term Delivery (Detailed)

Data Ingestion

1. Automate ingestion from authoritative sources into Redshift
   * Replace manual SAP BO extracts with automated pipelines.
   * Validate schemas, apply DQ checks, land into Redshift staging.
   * Ensures timeliness, accuracy, and auditability.
2. Capture adjustments and mappings via OCP portal
   * CFO UI/ROCP enforces controlled uploads with schema validation.
   * All uploads require stewardship approval and maintain full audit trails.
   * Provides transparency and governance for mappings/adjustments.
3. Tokenize sensitive data (Protegrity)
   * Tokenization at write prevents storage of cleartext PII/SPI.
   * Deterministic tokenization allows joins; non-deterministic secures high-sensitivity data.
   * JIT de-tokenization only for authorized roles.

Data Transformation

1. Replace EUDA (Excel) using TranslateX
   * Parse EUDA formulas/macros; convert to governed SQL.
   * Reviewed, optimized, and tested for parity.
   * Removes spreadsheet risk from critical workflows.
2. Replace Alteryx workflows using GitLab Duo
   * Convert Alteryx XML node flows into SQL pipelines.
   * SQL optimized for performance and transparency.
   * Reduces dependency on heavy, black-box Alteryx workflows.
3. Review, Optimize/fine tune the SQLs, Stitch the SQLs, Reconcile data between Alteryx output and SQL generated output, create views in Redshift and load the data into final tables.
   * *Review:* Peer review SQL for correctness, naming, null handling, surrogate keys.
   * *Optimize:* Tune joins, distribution/sort keys, remove inefficiencies.
   * *Stitch:* Consolidate EUDA/Alteryx-converted SQL into staging → mart pipeline.
   * *Reconcile:* Compare outputs to Alteryx golden results; correct discrepancies.
   * *Views in Redshift:* Create semantic views to abstract physical tables.
   * *Final Tables:* Load curated, partitioned query-ready tables.

Reporting

1. Point Tableau reports to Redshift final tables
   * Update workbook datasources to new semantic views.
   * Validate KPI parity, totals, and drill-down consistency.
   * Optimize live connections for interactivity.

Phase 1 Deliverables

* Automated ingestion pipelines into Redshift.
* Adjustments/mappings uploaded via OCP portal with audit trail.
* Tokenized Redshift datasets.
* SQL transformations replacing all EUDA/Alteryx logic.
* Certified semantic views and final Redshift tables.
* Tableau dashboards repointed and validated.
* Reconciliation packs documenting legacy vs new outputs.

Phase 1 Acceptance Criteria

* ≥99.9% reconciliation accuracy with legacy Alteryx/EUDA results.
* Pipelines run within SLA (T+1 7 AM ET).
* All sensitive data tokenized; zero cleartext findings.
* Tableau dashboards successfully cut over to Redshift.

8. Strategic Solution – Target Vision

* Config-driven ingestion pipelines into S3 Iceberg zones (Raw → Base → Prepared → Reporting).
* Transformations standardized using dbt, Airflow, Glue.
* Curated outputs stored in Iceberg; consumed by Snowflake and Databricks.
* Tableau and Power BI connected to certified marts.
* CFO Dial-In UI as the single launchpad across reporting tools.
* Governance: Lake Formation (catalog/ACLs), Alation (metadata/lineage), Immuta (row/column policies).
* JIT de-tokenization implemented in Snowflake/Databricks via secure UDFs.

9. Analysis Phase

* Source system analysis to identify authoritative feeds.
* Gap analysis: map EUDA and Alteryx logic to SQL equivalents.
* Define DQ and reconciliation rules.
* Assess tokenization policies for classification accuracy.

10. Development Phase

* Build pipelines in Glue/Airflow with tokenization at ingestion.
* Convert EUDA and Alteryx workflows into optimized SQL.
* Implement dbt models (staging, intermediate, marts).
* Deploy Iceberg tables with partitioning and schema evolution.
* Register datasets in Lake Formation; sync metadata/lineage in Alation.
* Build CFO UI modules for uploads, writeback, and monitoring.

11. Testing Phase

* Unit testing: SQL correctness, schema compliance.
* Integration testing: Ingestion → transformation → reporting.
* Reconciliation testing: Golden dataset comparisons, ≤0.1% variance.
* Performance testing: Pipeline SLA and BI interactivity targets.
* Security testing: Verify masking, JIT de-tokenization, pen testing.
* UAT: Finance validation of dashboards.

12. R&D / POC

* Validate TranslateX conversion accuracy for EUDAs.
* Validate GitLab Duo conversion coverage for Alteryx XML.
* POC JIT de-tokenization in Snowflake & Databricks.
* Validate Power BI connectivity and performance.

13. Risks & Mitigations

* Conversion gaps (EUDA/Alteryx): Mitigation → reconciliation & manual SQL refinement.
* Security breaches: Mitigation → tokenization + Immuta default masking, audits.
* User adoption: Mitigation → training, phased rollout, CFO UI simplification.
* Performance bottlenecks: Mitigation → SQL tuning, caching, scalable infra.

14. Deliverables

* Phase 1 pipelines, Redshift final tables, and Tableau cutover.
* Strategic onboarding of dbt, Airflow, Snowflake, Databricks, Power BI.
* CFO Dial-In UI enhancements.
* Governance integrations: Lake Formation, Alation, Immuta.
* Reconciliation and validation packs.
* Decommissioned EUDA and Alteryx workflows.

15. Acceptance Criteria

* Phase 1 deliverables signed off.
* Strategic components onboarded and operational.
* 100% in-scope reports automated.
* Compliance with all audit/security requirements.
* CFO UI operational as the single reporting launchpad.

16. Governance

* Steering Committee: Project direction and risk oversight.
* Data Governance Council: Stewardship, lineage, data quality enforcement.
* Change Control Board: Approves enhancements/scope changes.
* PMO Reporting: Weekly status, risks, dependencies, and metrics.