

AccionLabs - Enterprise Data Transformation Services





Agenda



- About Accion Labs
- Enterprise Data Capabilities Overview
- Data Strategy
- Data Governance
- Realtime Analytics
- GenAI Capabilities
- Partnerships
- Databricks Capabilities
- End to End Solution Case Studies

About Accion Labs

5500+

Global Team
Members

1500+

Projects
Delivered

25+

IP
Accelerators

170+

Clients
(Long term relation)

INDUSTRY CERTIFIED

ISO
9001:2015

ISO
27001:2013

ISO
27701:2019

CMMI
V2.0 Level3

PCI DSS
Compliant

HIPAA
Compliant

TECHNOLOGY PARTNERS



servicenow



Adobe Commerce Cloud



HOW WE WORK

Data at Heart

Accion Labs reimagines the possibilities for digital transformation from innovation to growth

Digital in Thinking

Scalability, manageability, robustness, and efficiency are always their top priorities

Innovation in Approach

Everything is planned to minute detail, ensuring high-quality deliveries to clients every time



OUR GROUP COMPANIES

Local Partner - Global Flavours



5500+
Employees

100+
Clients

50+
Tools/IP

21
Global
Offices

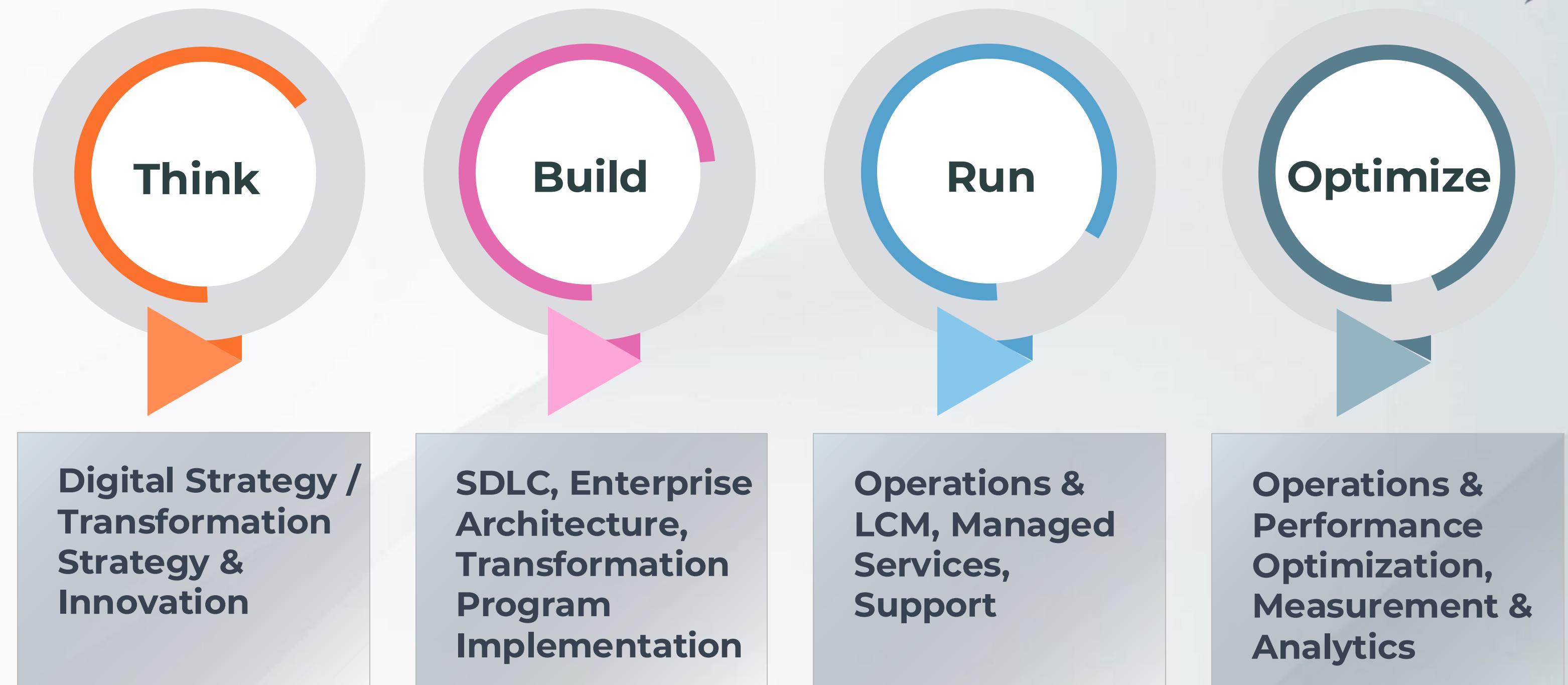
14
Development
Centers



Our Services



Drive rapid business growth through technology



Enterprise Solutions & Consulting

Enterprise Transformation
(Technology)

Managed Services -
Business Applications and Products

Enterprise Process Implementations & Optimizations using Platforms

Legacy Modernization and Re-engineering

Enterprise Data Transformation Strategy

Product Portfolio Rationalization

Data & Analytics Strategy | AI Strategy

Professional Services

Our Enterprise Transformation Services



Process Digitization & Optimization

- Biz Process Implementation
- Business process Customization & Optimization
- Workflow based applications
- Integration Platforms | Enterprise Integrations
- Robotic Process Automation
- Gen-AI based Process Automation

Tech Landscape Rationalization

- Enterprise Architecture
- Application portfolio Rationalization
- Re-engineering & Modernization
- Product Engineering
- Design Engineering | UX / UI
- IoT Applications

Data Transformation

- Org-wide Data Strategy | Data Model
- Data Governance & Data Quality
- Data Modernization | Data Migration
- Data Engineering | MLOps, AI Ops, Data Ops
- Analytics – Bi. Reporting, Realtime | AI/ML
- GenAI Powered Analytics & Insights

Cloud Transformation

- Org Cloud Strategy & Roadmap
- Application and Infra Cloud Migration
- GenAI-powered Managed Services
- CloudOps | DevOps | DevSecOps
- Cloud Native Business Apps
- Cloud Platform & Infra Services

MS Dynamics

Salesforce

ServiceNow

LC / NC Platforms - Power Platform, Mulesoft, Boomi, Appian

RPA

.NET, JAVA, Python

Magento

AEM

Snowflake

Azure Synapse & Fabric

AWS Redshift

Databricks

Azure

AWS

GCP

Cloud Platform & Data Services

Gen AI Powered (*Potentially*)

Microsoft Azure Cloud Services | Azure Process Automation | Azure Data Services

AWS Cloud Services, Platform Services, Data Services



Partnerships and Alliances

servicenow.

 **snowflake**

 **Microsoft**



Adobe Commerce Cloud

 **amazon web services™**

 **databricks**



 **Google**
Cloud Platform

 **HORTONWORKS®**

 **mongoDB**

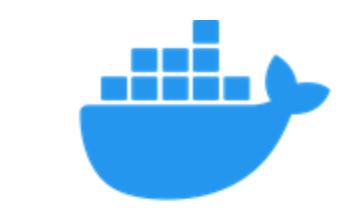
 **calm.io**

 **kony**

 **CHEF**



 **UiPath™**

 **docker**

 **MariaDB®**

Enterprise Data Capabilities - Overview



Service Segments

Focus Areas

Corner Stones

Data Science and AI Services

Strategy and Roadmap
(Assess – Guidance – Maturity)

AI & ML Model
(Develop – Evaluate –Deploy)

Productization and Management
(Accelerate – Operationalize – Democratize)

Business Advisory

Data Discovery and Data Mining

ML Modeling, Training, EDA

MLOps and Scaling AI

AI & ML Platforms

Data Sharing

Data Storytelling

Technology Stack

KPIs and Metrics

NLP, NLG, NLQ,

Responsible AI, GenAI, Trusted AI

CoEs, Workshops

GenAI, GenBI and Self-serve

Optimal Dashboards

Data and Analytics Modernization Services

Implementation and Integration
(Design-Build-Integrate)

Migration and Modernization
(Run-Optimize-Scale)

Data Management & Governance
(Policy – Regulation – Execution)

Data Engineering

Pipeline Automation BI & Reporting

Hyperscalers, Data Clouds & Platforms

DataOps and FinOps

Data Lineage, Cataloging, Data Marketplace

Data Trust, Security, Compliance & Access

Data Fabric and Mesh

Data Hub, Streaming Data

EDW, Lake House

Edge & Distributed Computing

Data Quality & Observability

MDM, Metadata and Data Ecosystems

Tools & Accelerators

Industry Domain Expertise

Experience Engagement

Governance Ethics

Co-innovation Collaboration

Change Management



Enterprise Data Capabilities

Enterprise-wide Data Transformation Solutioning



ENTERPRISE-WIDE DATA STRATEGY

- Business Aligned Enterprise-wide Data Strategy
- Application Data Model
- Master Data Strategy & MDM Solutioning
- Continuous Data Strategy Assessment and realignment
- Continuous Monitoring and Improvement

DATA GOVERNANCE

- Enterprise-wide Data Governance Strategy
- Governance Policies, Frameworks, Guidelines
- Data Catalog, Data Lineage
- Data Governance Tooling

DATA QUALITY

- Identify and Establish Enterprise-wide Data Quality Standards
- Solutions to push Data Quality standards across all applications dealing with data
- Data Quality Tooling Solutions – Data at Rest + Data in Transit
- Data Quality Monitoring

DATA MANAGEMENT

- Defining & Establishing Data Management Guidelines
- Data Architecture
- Data Engineering
- Data Analytics

DATA INTEGRATION / EXCHANGE

- Defining Data Integration & Exchange Strategy and Mechanisms
- Defining & Creating Data Integration Policies

DATA MONETIZATION

- Defining scope of Data-Driven for Enterprises
- Setting up Data Landscape for Data-Driven Decision-making

Data & Analytics Solutioning



DATA & ANALYTICS ROADMAP

- Data Platforms & Analytics Landscape Assessment
- Data & Analytics Strategy Creation
- Long Term Data & Analytics Roadmap

DATA LANDSCAPE MODERNIZATION

- Composable Data Landscape & Architecture
- SSOT / MSOT
- Low-Code / No-Code
- XOps Enabled (DataOps, MLOps, AIOps, GenAIOps)
- Data as-a-Service
- Analytics as-a-Service

BI LANDSCAPE CONSOLIDATION & MODERNIZATION

- Migration and Consolidation of diverse BI tools landscape into a leaner BI landscape
- Self-Service Dashboards

ANALYTICS: BI / REPORTING / ML / AI

- BI and Reporting | Automated Reports & Dashboards
- ML Modeling
- Predictive Analytics

LEGACY DATA PLATFORMS CONSOLIDATION & MIGRATIONS

- Migration of Legacy Data Systems to New platforms
- Migration to Cloud-based Managed Data Platforms
- Consolidation of Multiple Data Warehouses into Unified Data Platforms

GEN-AI POWERED DATA ANALYTICS

- Knowledge mining & Content Creation
- Conversational Data Analytics
- Process Optimization

Partnerships & Technology Spectrum



Accion Partnerships

Microsoft
Partner

Gold Cloud Platform
Gold Application Development
Silver DevOps



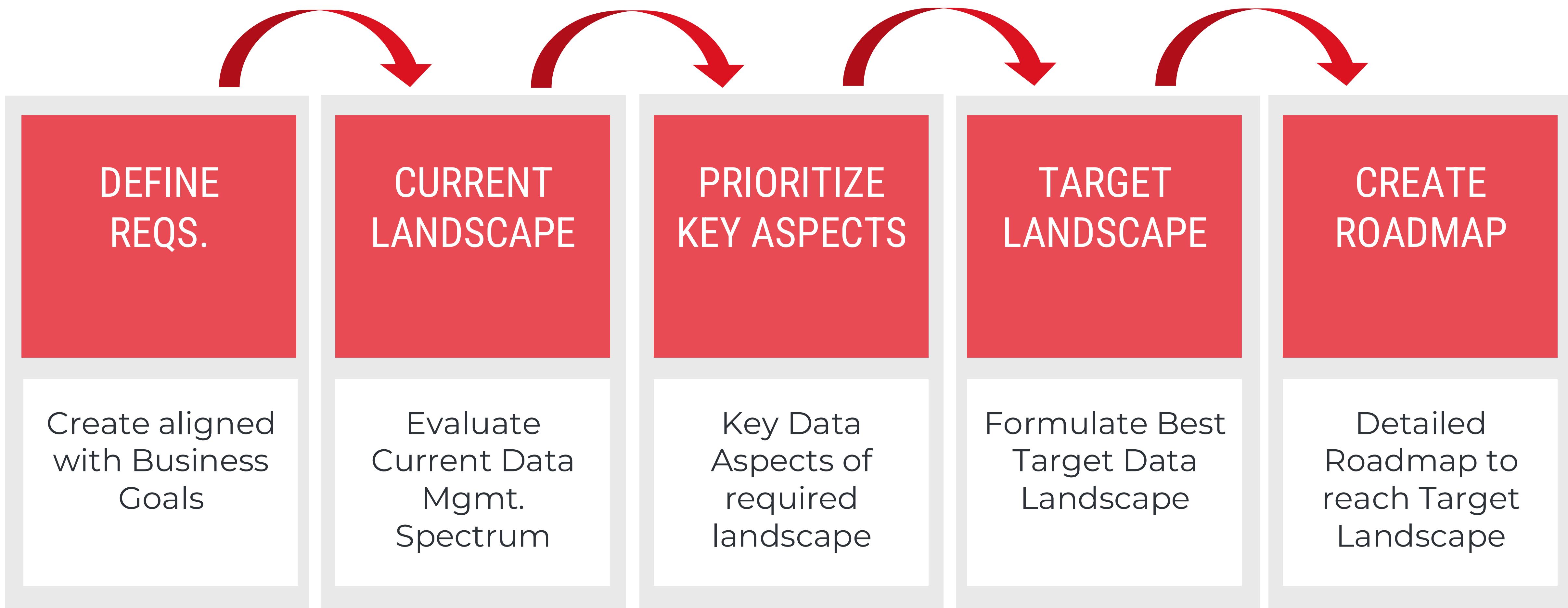
The logo for Amazon Web Services (AWS) features three orange 3D cubes arranged in a staggered, upward-pointing triangle. To the right of the cubes, the word "amazon" is written in a lowercase, bold, black sans-serif font. Below "amazon", the words "web services" are written in a smaller, lowercase, black sans-serif font.

Partner
Network

We work actively for clients in 120+ open-source and commercial distributions, application platforms, and data stores including



Data Transformation Approach (Business - Ops - IT)





DataLabs:

Accion's Enterprise Data & Analytics Landscape Maturity Assessment Program

Context: Enterprise Data & Analytics Ownership & Challenges



1 BUSINESS

Business Functions (e.g., Finance, Mktg., Commercial, Mfg., R&D, HR)

- | Challenges |
|--|
| <ul style="list-style-type: none">▪ Lack of trust on Data▪ Unavailability of Data at the right time in right form▪ Data Reconciliation▪ Late and Incorrect Insights▪ Repetitive and Manual Reporting▪ Fragmented, Siloed & Duplicated Analytics |

- | Impact |
|---|
| <ul style="list-style-type: none">○ Slow and Manual decision making○ Up to 70% higher efforts due to data credibility and reconciliation needs |

Data & Analytics Ownership & Challenges

2 OPERATIONS

Operations and Change Mgmt.
(Group that works closely with Business & IT)

- | Challenges |
|--|
| <ul style="list-style-type: none">▪ Lack of Ownership Matrix▪ Lack of Data Classification▪ Hard to search Data▪ Lack of Governance & Data Mgmt.▪ Lack of DataOps & MLOps▪ Inefficient Change Mgmt.▪ Manual Data Reconciliation |

- | Impact |
|--|
| <ul style="list-style-type: none">○ 50% higher costs in Operations○ 30%-40% higher cost in Data Management & Governance |

3 IT / ENGINEERING

Technology
(Engineering & Infrastructure)

- | Challenges |
|---|
| <ul style="list-style-type: none">▪ Lack of Data Warehouse and models▪ Fragmented data modeling▪ Non-scalable Architecture▪ Inconsistent Data Quality, Duplicate Data▪ Lack of Data Lineage |

- | Impact |
|---|
| <ul style="list-style-type: none">○ Up to 40% higher cost due to fragmented data landscape, Infrastructure & Tools |

DataLabs – Data Landscape Maturity Assessment



- ✓ Evaluates Adherence of Data Landscape to Industry Standard via **450** Criteria points

Measures adherence to Industry best practices via an Exhaustive Criteria List

1

450 Criteria Points for
Industry standard

from

70 Data
Categories

from

8 Core
Areas

across

3 Buckets

Business
IT & Engineering
Operations

2

Evaluate:

- ✓ **DATA LANDSCAPE MATURITY SCORE (0-5)**
- ✓ **MATURITY POLYGON**

Maturity Score (0-5)

4 - 5

ADVANCED

3 - 4

PROGRESSIVE

2 - 3

BASIC

1 - 2

REACTIVE

0 - 1

PRIMITIVE

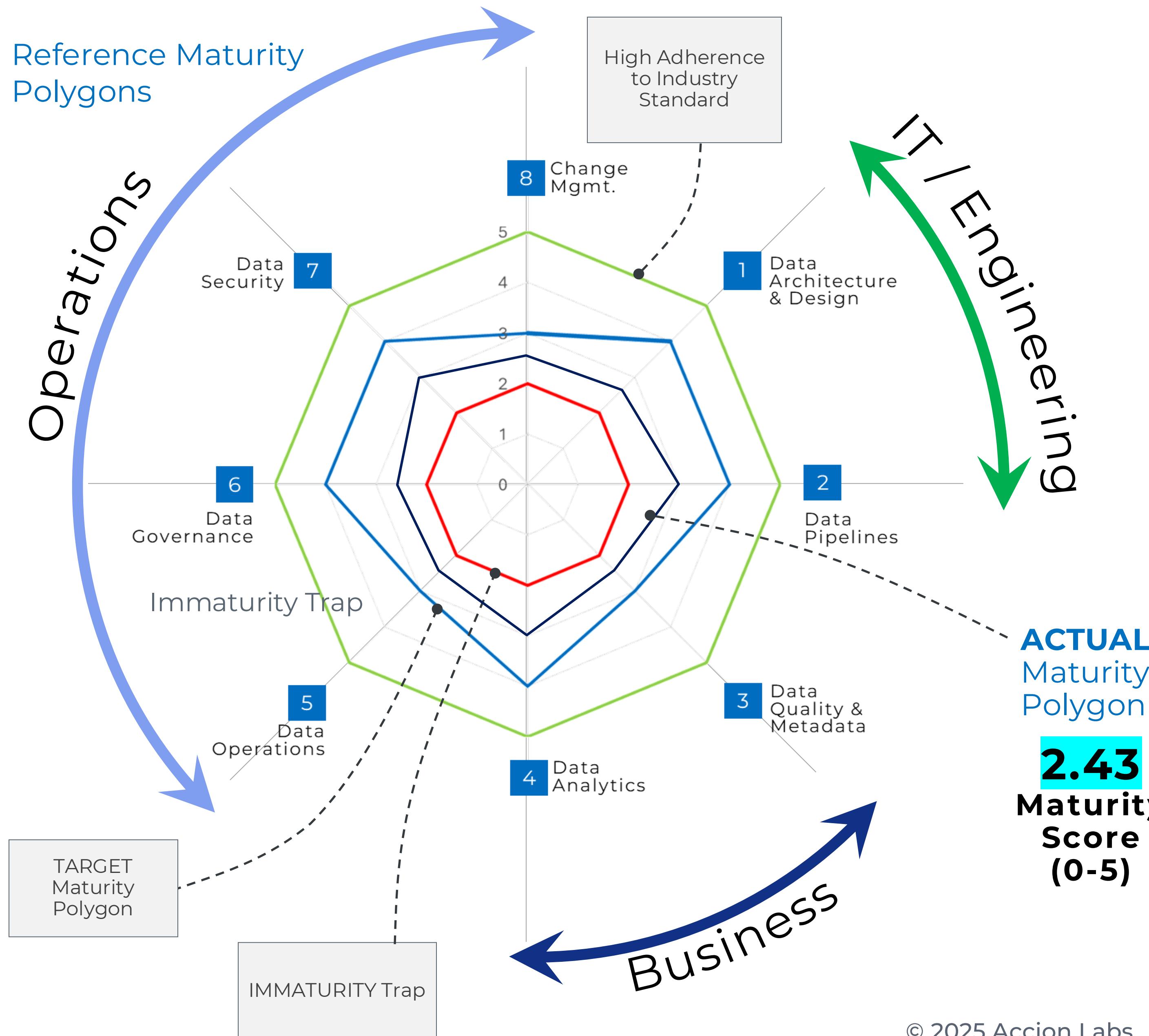
3

Provide:

- ✓ **TARGET MATURITY**
- ✓ **TARGET LANDSCAPE**
- ✓ **DETAILED ROADMAP & RECOMMENDATIONS**

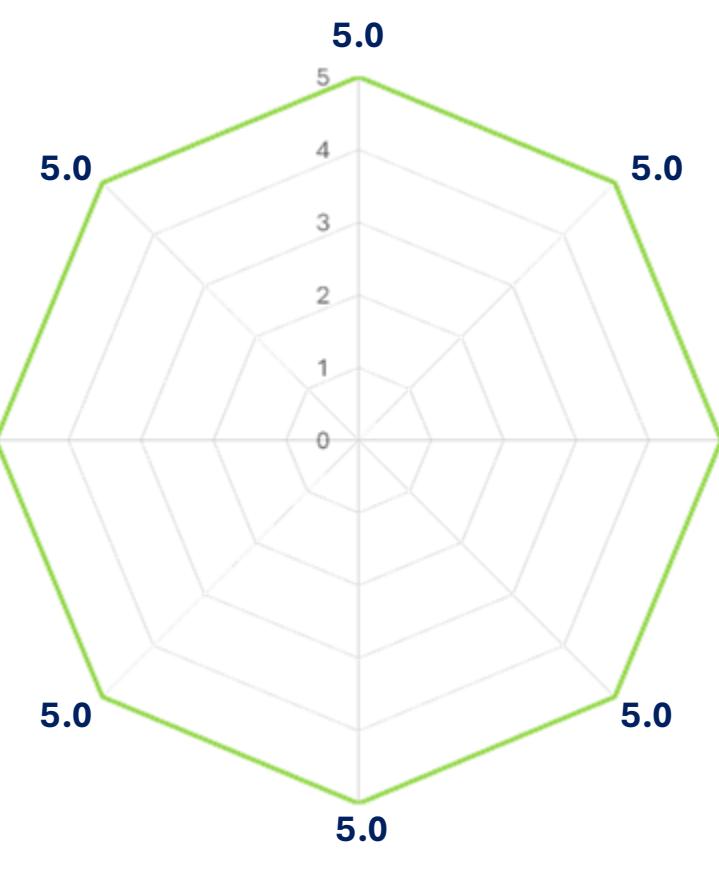
to reach the Target Maturity & Target landscape

DataLabs – Maturity Polygon & Maturity Score

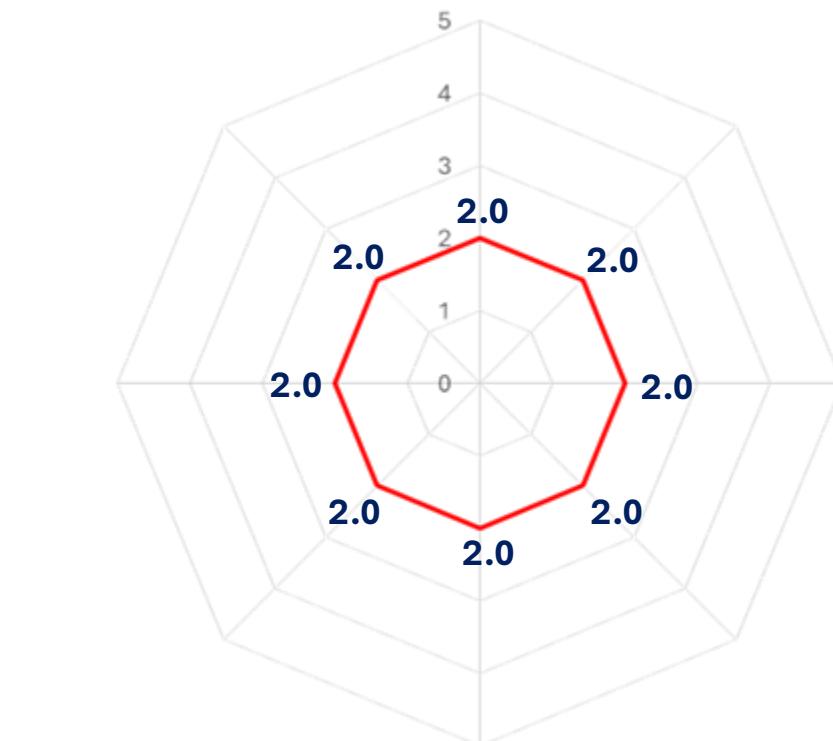


ACTUAL
Maturity
Polygon

2.43
Maturity
Score
(0-5)



Example Polygon
With
Score of 2.43



Assessment Model



Get your Data Landscape evaluated via one of the 3 PROGRAMS below to get Data Landscape Maturity score, along with concrete findings, impact & recommendations to reach TARGET score & Landscape.

STARTER

1 Week

Data Landscape Maturity Score of **One Group** with Key Observations

- Assessment on all Areas in any one of the following buckets:
 - Business
 - IT/Engineering
 - Operational
- High-level Findings
- Brief Recommendations

STANDARD

2 Weeks

Overall High-level Data Landscape Maturity Assessment With high level Recommendations

- Assessment on all Data Functions and areas
- Engagement with key IT leaders
- Evaluation of a Maturity Score for overall Data landscape as well as individual data functions
- Brief Recommendations

ADVANCED

Recommended | 4 weeks

Overall Data Landscape Maturity Score + Detailed **Analysis** + Concrete Short-term & long-term **Recommendations**

- Detailed Assessment on all data areas
- Deep-dive analysis where we engage with key Business and IT leaders of that function for 20 days
- Maturity Scores for overall Data landscape as well as individual data areas
- Concrete recommendations with roadmap to advance to Target maturity level

Maturity Assessment Steps



1

Stakeholder Discussions

- Discussion with Business Stakeholders.
- Discussions with Technology Stakeholders.
- Gathering and compilation of responses and answers on the *Maturity Evaluation criteria*.

2

Defining Applicability

- Alignment of Areas and Criteria of evaluation based on *Business Priorities and Relevance*.
- Enabling/Disabling of applicable/inapplicable Criteria.

3

Assigning Weightages

- Assignment of Weightage to each applicable group (*Total weightage* across all 'applicable groups' to be =100%).
- Assign Weightage to and each applicable category within each applicable group (*Total weightage* across all 'applicable categories' within the group to be =100%).

4

Maturity Score

- Evaluation of Group-wise Maturity score (Score between 0 to 5).
- Evaluation Overall Maturity score (Score between 0 to 5).

5

Detailed Recommendations

- Recommendation of Target Maturity.
- Detailed recommendations on overall landscape and for advancing maturity score to Target Maturity.
- Detailed recommendations at each aspect (group) in order to advance the maturity score to Target Maturity.

6

Implementation Roadmap

- Discussion with Business Stakeholders.
- Discussions with Technology Stakeholders.
- Gather and compile answers on Evaluation criteria.

Strategy Assessment Deep Dive



1 BUSINESS

Data Analytics

- Analytics modelling ? Data aggregates?
- Shareable and discoverable?
- Self-Serving reports, Dynamic, DIY?
- Data Science, Predictive ML/AI?

Data Quality | Metadata Quality

- Data Quality, Availability, Integrity, usability, security – In-flight and At-Rest
- Classification - Business Vs. Technical
- Data Cataloging, Data Discoverability, Searchability, Change history, Audit trail

2 OPERATIONS

Data Operations

- Engineering Operations
- Program & Project Mgmt.
- Build Mgmt. and CI/CD
- Business SLAs

Data Security

- Classifying Sensitive Data
- Data Level Security
- App Security
- Infra Security
- Security Best practices

3 IT / TECH

Data Strategy

- Information Architecture strategy
- Business Alignment

Data Architecture & Design

- Scalable architecture?
- Expandable? Mutable?
- Provider of only data or also Analytics?
- Modelling Maturity of core business entities and transactions?

Data Pipelining

- ETL Scalable? Reusable? Automated?
- ETL design, Data Quality, Deduplication, Error handling, Alerting, Caching, Data Recovery, Zero Loss

Change Management

- Standard Change Mgmt. Operational Model
- Common Issue Mgmt. system
- Observability

Data Governance

- Ownership between Biz & Tech
- Roles and Responsibilities
- Data Access Policies
- Regulatory Compliance
- Data Management

Mapping of Data Functions to Assessment Buckets



3 Data functions mapped to
3 Assessment Buckets



Data Landscape evaluation
8 Core Areas | 70 Categories |
450+ Criteria Points

Core Areas

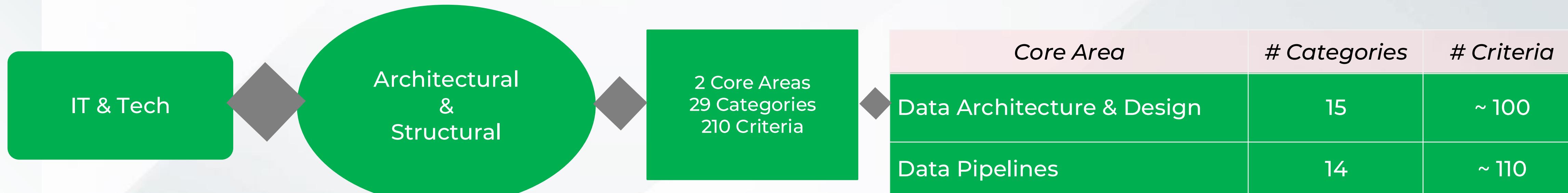
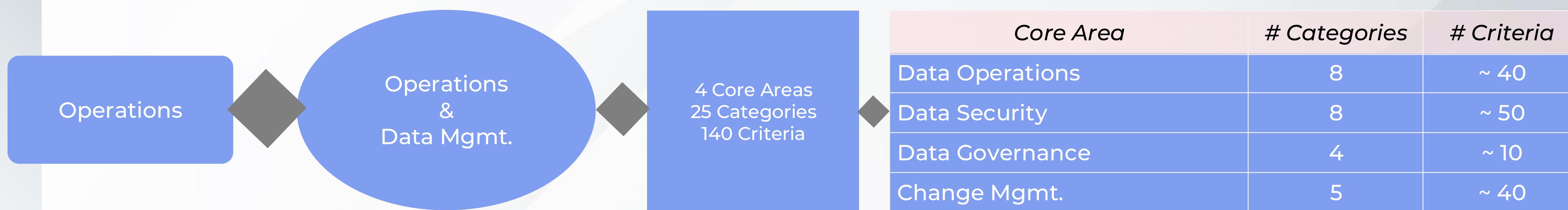
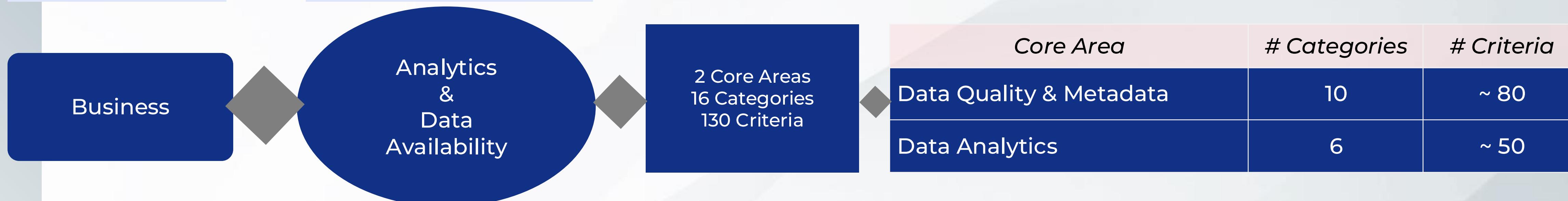
8

Categories

70

Criteria

450+



Result of Maturity Assessment



Maturity Score (0-5) along each of the 8 Axes (8 Core areas)

Maturity Score (0-5) for the entire landscape

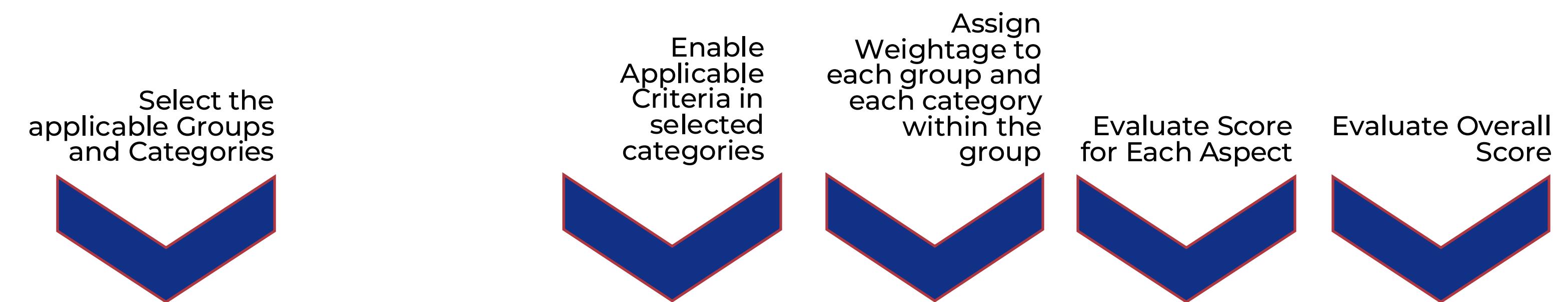
Maturity Polygon that depicts the landscape's adherence to Industry Best Practices in each ownership area – **BUSINESS**, **IT/Engineering** and **Operations**, and in each of the 8 key areas

Individual and Overall Maturity Scores



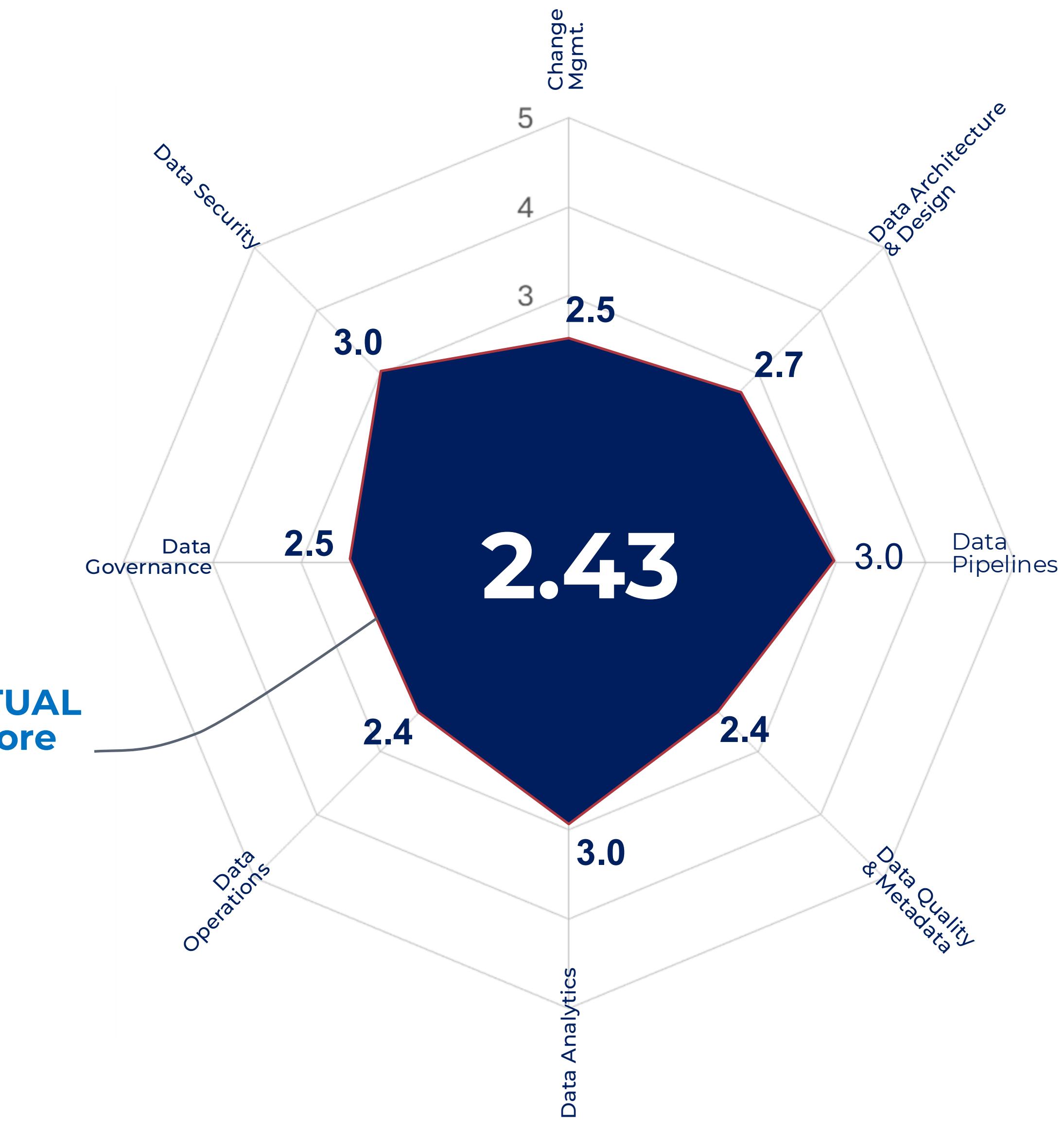
Example:

Defining Applicability,
weightages and arriving at
MATURITY SCORES

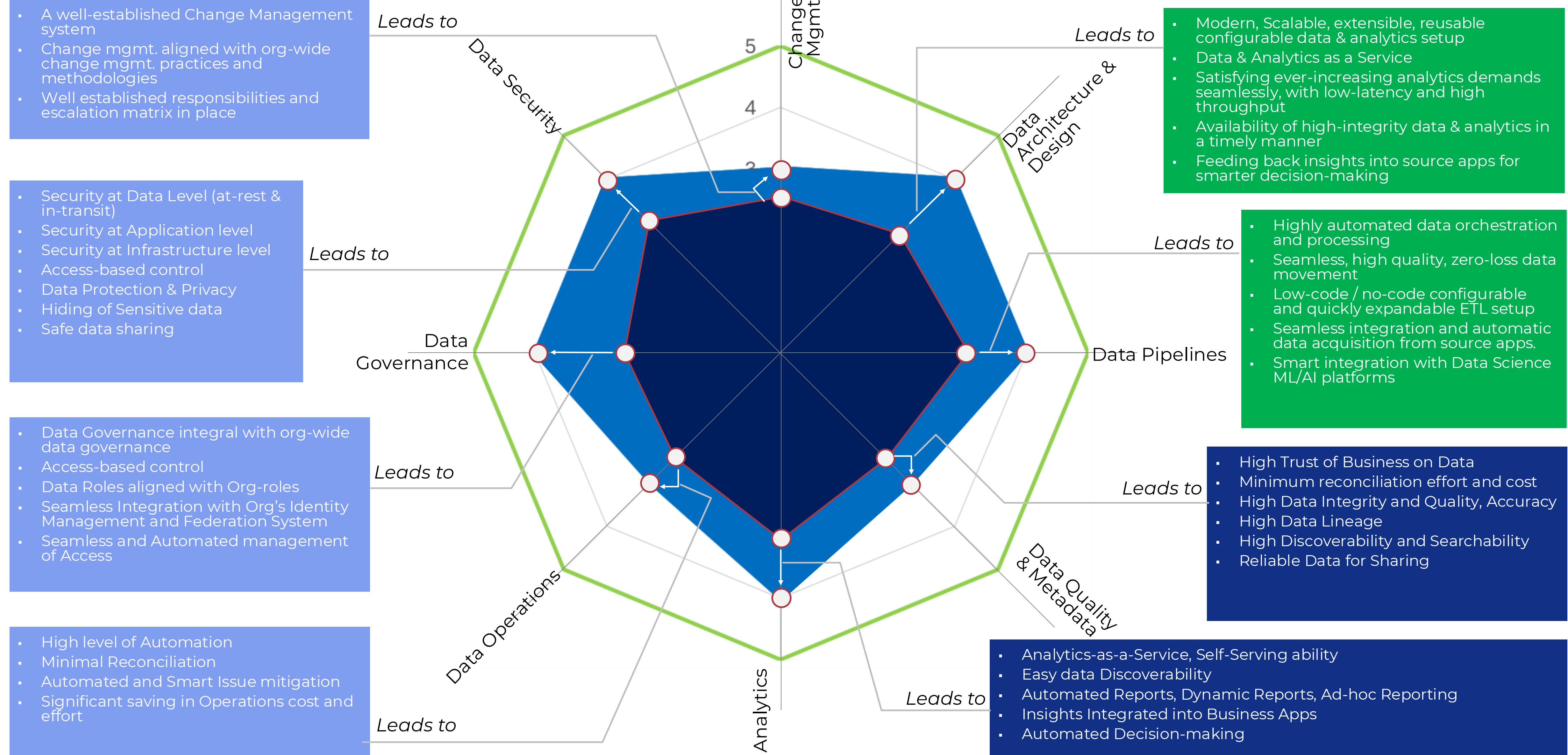


ASPECT (Area)	Total Categories	Enabled (Applicable) Categories	Total Criteria points	Enabled (Applicable) Criteria Points	Assigned Weightage (Total 100%)	Score (1 to 5)	Score (1 to 5)
Data Analytics	10	6	~ 80	~ 40	10%	2.75	2.43
Data Quality, Metadata	6	9	~ 50	~ 70	15%	2.23	
Data Security	8	6	~ 40	~ 30	8%	2.55	
Data Governance	7	5	~ 50	~ 30	8%	2.17	
Data Ops	4	4	~ 10	~ 10	6%	2.80	
Change Management	5	4	~ 40	~ 35	3%	2.60	
Data Architecture & Design	15	15	~ 100	~ 100	20%	2.56	
Data Pipelines	14	12	~ 110	~ 85	15%	3.62	

Maturity Polygon - Comparison with Industry Standards



Detailed Recommendations for Moving To Target Maturity



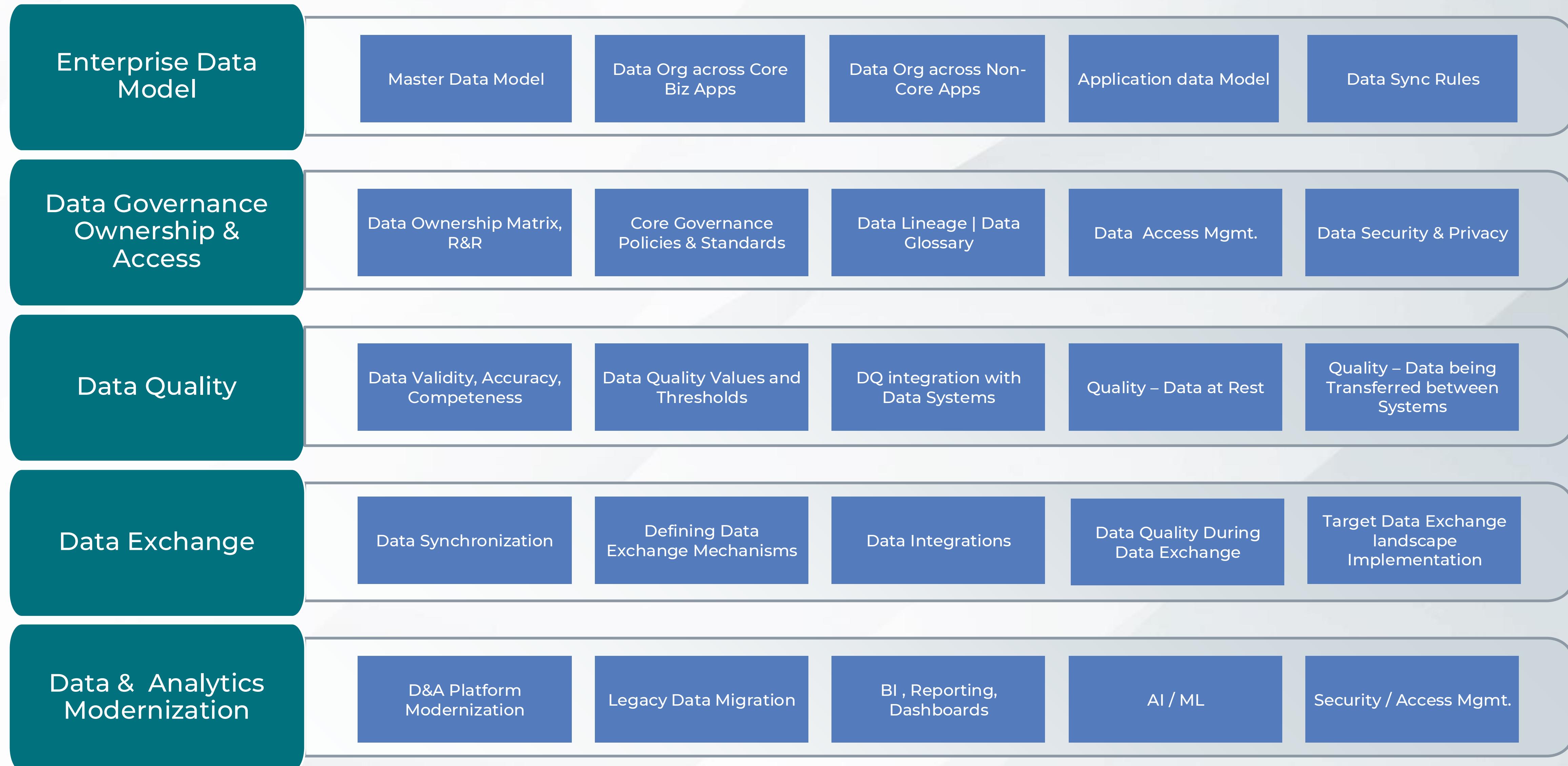


Data Governance



- ▶ Data Governance Strategy
- ▶ MDM Approach
- ▶ Case studies

Accion's Focus - Enterprise Data Management Spectrum



Data Governance Related Services



Services

DATA CATALOG

DATA QUALITY & INTEGRITY

DATA LINEAGE

DATA DISCOVERY & SEARCH

MASTER DATA MGMT.

DATA PRIVACY

DATA SECURITY

DATA COLLABORATION

DATA GOVERNANCE ROADMAP

DATA RULES, POLICIES & PROCESSES

STAKEHOLDER MGMT.

Tools

 Alation

 Collibra

 Microsoft Purview

 atlan

 Informatica

 AWS Glue Catalog

 AWS Lake Formation

 GoodData

 ataccama



Data Governance Approach



1. Select an Approach for Implementing Data Governance



Top-down

1. Bring All Business Owners Onboard
▪ Org-wide Governance Body
2. Create Common Master Data Definition (All MDM entities)
3. Create Comprehensive MDM Strategy
▪ Create Comprehensive MDM Solution
▪ Create Common MDM Hub
4. Update all Business Systems with New MD Model
5. Establish Exhaustive Data Quality
6. Update all Biz App Workflows
▪ Exhaustive Testing
7. Fix All Inter-system Data Exchanges, Data Synchronizations, Data Interfaces
8. Create Modern Data & Analytics Platform
▪ Rebuild Data Analytics / Reporting

Hybrid

Bottom-up

7. BI Reporting Consolidations
6. Data Platforms / data Warehouses Consolidation
5. Create a Common Data Master
▪ Reconcile Ingested Data with the Master
4. Identify Data Reconciliation Problems due to different Versions of Master Data
3. Analyze Current Data Platforms and BI tools (per Business Function)
2. Identify Business KPIs
▪ Identify Source Systems for All KPIs
1. Identify Consumers of Analytics
▪ (Business Functions)

Guided Top-down

- Identify Core MDM Entities
- Identify systems of Highest Impact
- Identify Owners
- Limited Master Data Fix
- Optimum Data Governance
- Optimum Data Quality Around Selected Data
- Data & Analytics Modernization

Toward Target Strategy

2. Assess Current Data Landscape



Evaluate the application ecosystem to find out the core landscape that deals with **CUSTOMER** and **PRODUCT**



Identify **AREAS OF DATA RECONCILIATION** for the core applications set, and the business impact of those reconciliations



Identify **DATA EXCHANGE SCENARIOS** and situation for exchange of Customer and Products data between applications

Establish the "**CORE APPLICATIONS**" set - the set of applications and their dependencies which our assessment will focus on



IDENTIFY IMPACTED APPLICATIONS AND PROCESSES impacted by Customers and Subscriptions



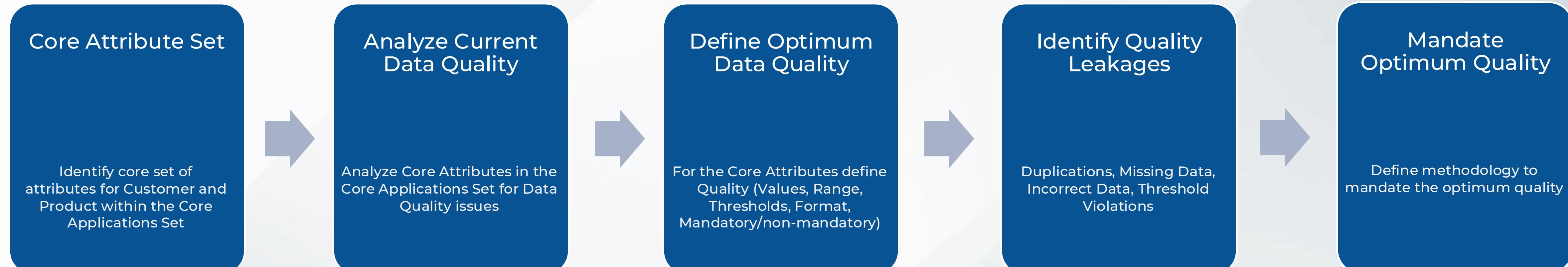
3. Identify Initial Problem statement (“Start Line”)



Track 1: Establishing core Governance group and framework



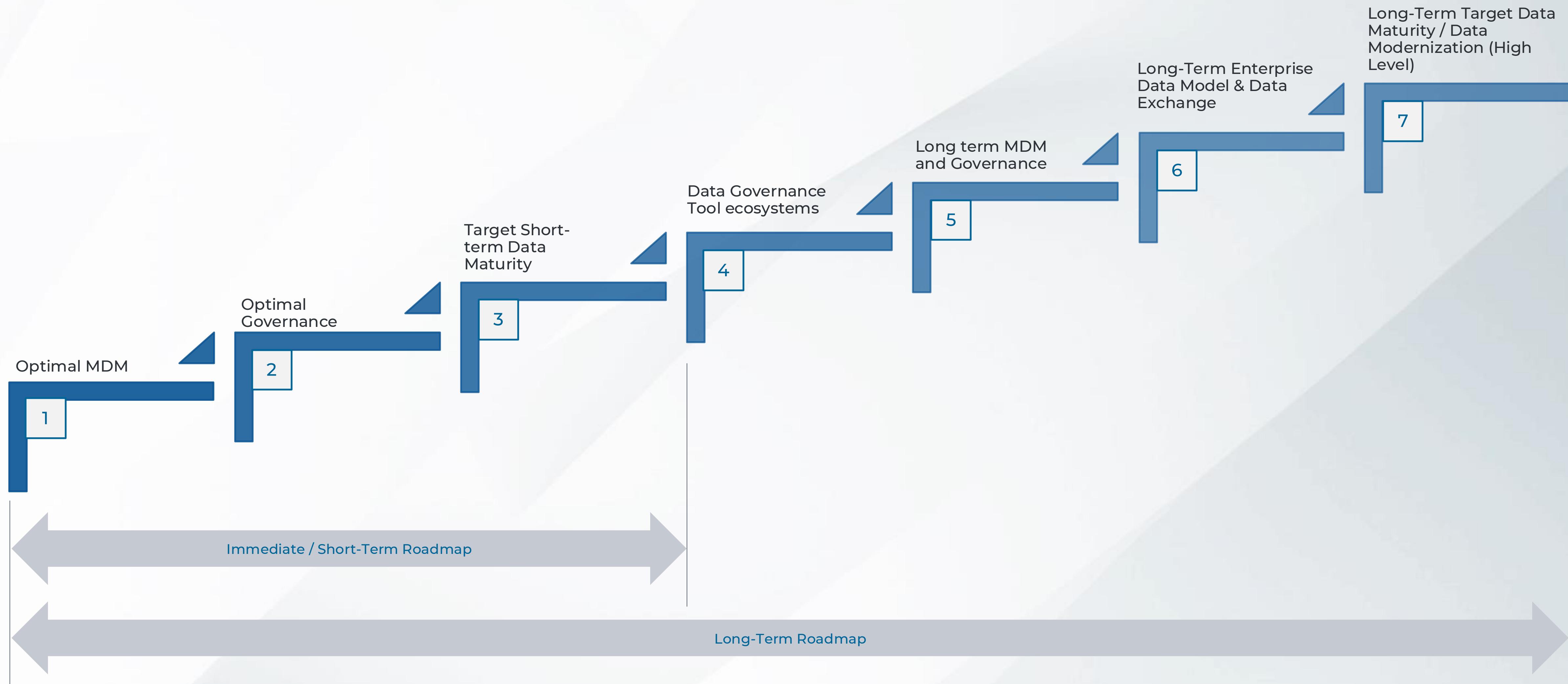
Track 2: Identifying Initial level of Data Quality & Data Governance



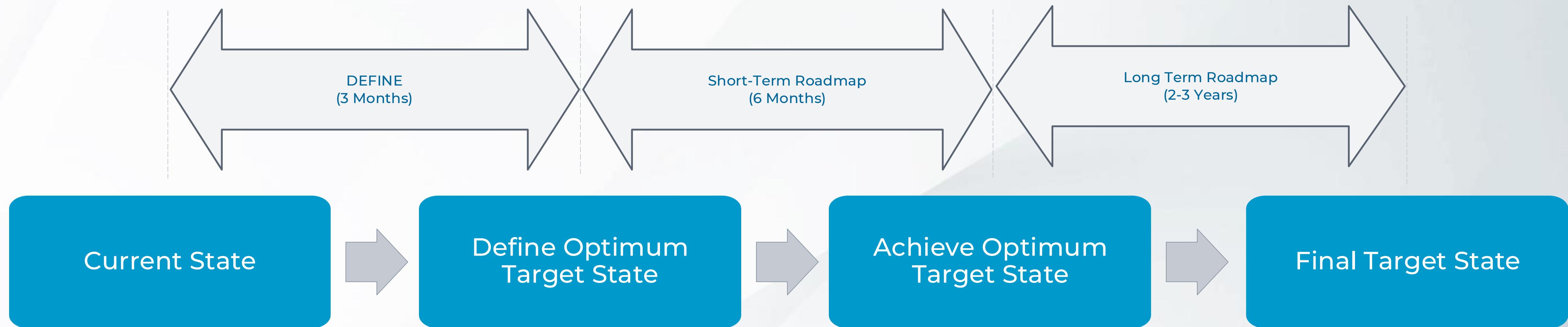
4. Define road to reach Intermediate Maturity State



5. Reaching Target State (Target Data Landscape Maturity)



Recap - The Data Landscape Maturity Journey



ASSESS

- Current Master Data Model
- Current Data Quality
- Current Data Ownership and Autonomy
- Current Data & Analytics Landscape

DEFINE

- Optimum Master Data Model
- Optimum Data Quality
- Optimum Data Governance framework
- Optimum Data & Analytics Landscape

ESTABLISH

- Optimum Master Data Model
- Optimum Data Quality
- Optimum Data Governance framework
- Optimum Data & Analytics Landscape

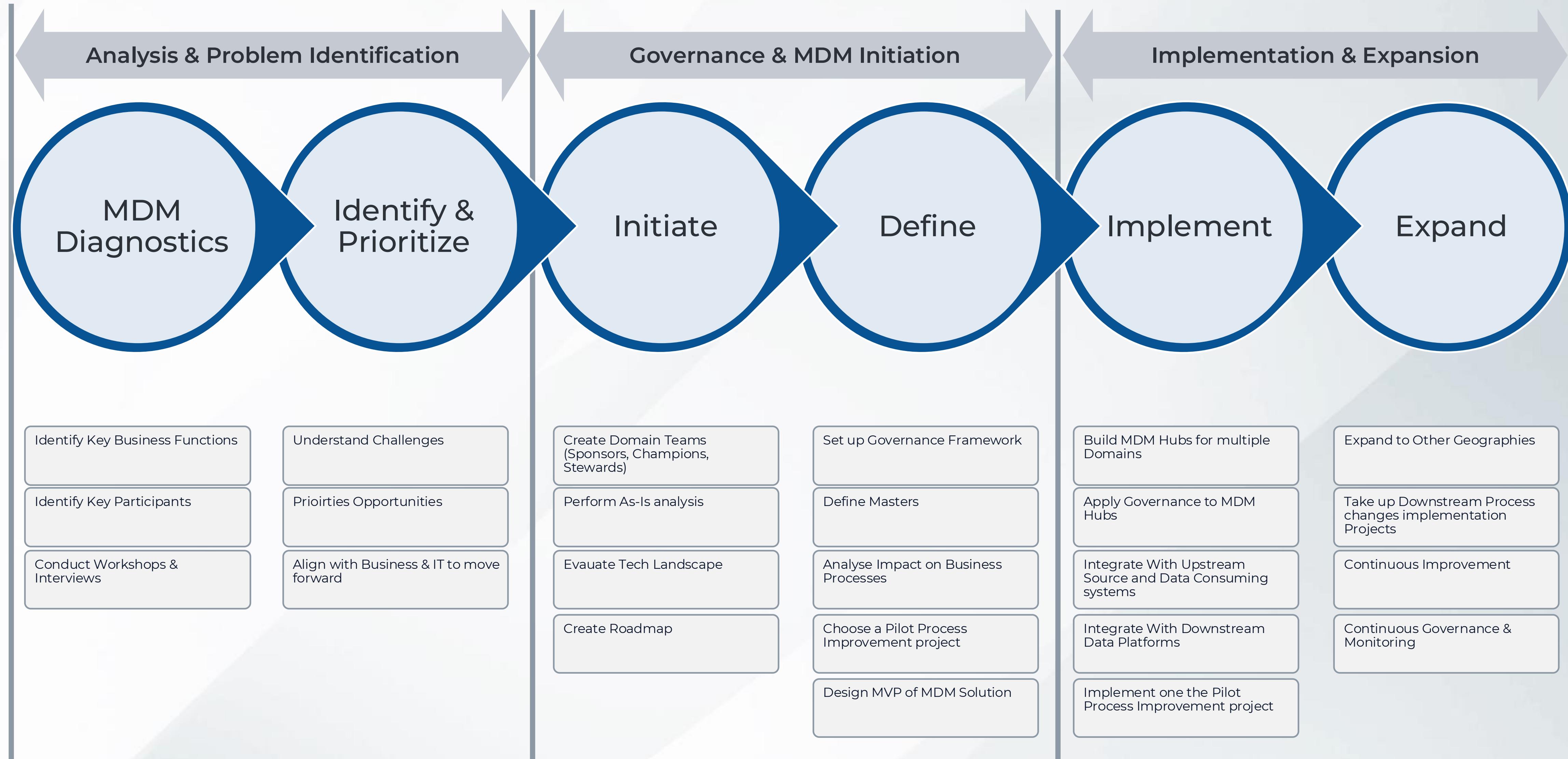
EXTEND

- Full Master Data Model
- Full Data Quality
- Full Data Governance framework
- Modernized Data & Analytics Landscape



MDM Transformation Approach

MDM Transformation Approach



Our Approach on MDM Maturity Enhancement



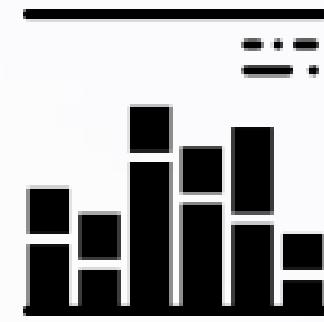
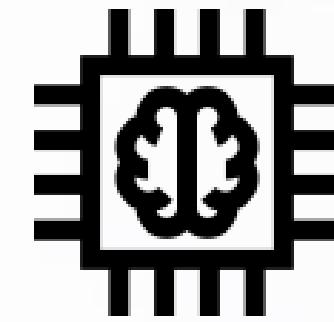


Our Process for Success

1. Integrated program management with CxO-level oversight
2. Expert teams with proven innovation
3. Deliverables with acceptance tests and full specifications
4. Focus on outcomes with tools for client independence



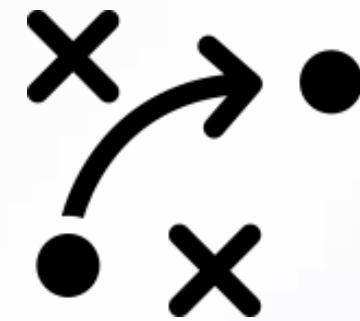
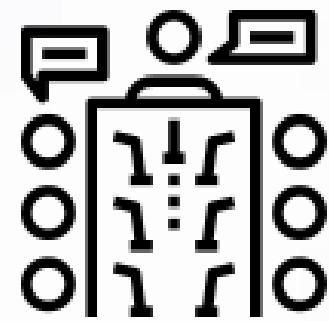
How We Strive for Client Success



Experts “On the Ground”



Integrated Program Management



CxO-level Management Oversight and Coaching



Engagements ranging from “Expertise” to “Owned Outcome”



“Sponsor-Aligned, Stakeholder-Ready” Updates



Data Governance Case Studies

Consolidation of multiple OpCos (1/2)



Company

A multi-billion USD Loyalty management company managing Loyalty points for 6 Airlines (International + Domestic) from multiple countries.

Extended Loyalty program to 30+ Partner companies

Challenge

- Data Contracts with 6 OpCos and 30+ Partners (Partner system continuously expanding).
- 4 Disparate Data systems handled Data & Analytics requirement for Loyalty Mgmt. Company.
- Data Governance required common tools across different technologies-based data platforms.
- Significant collaboration required with OpCos on the upstream side and partners on the downstream end..

OpCos

6
OpCos

12+
OpCo Sources

Loyalty Mgmt.

35+
Applications

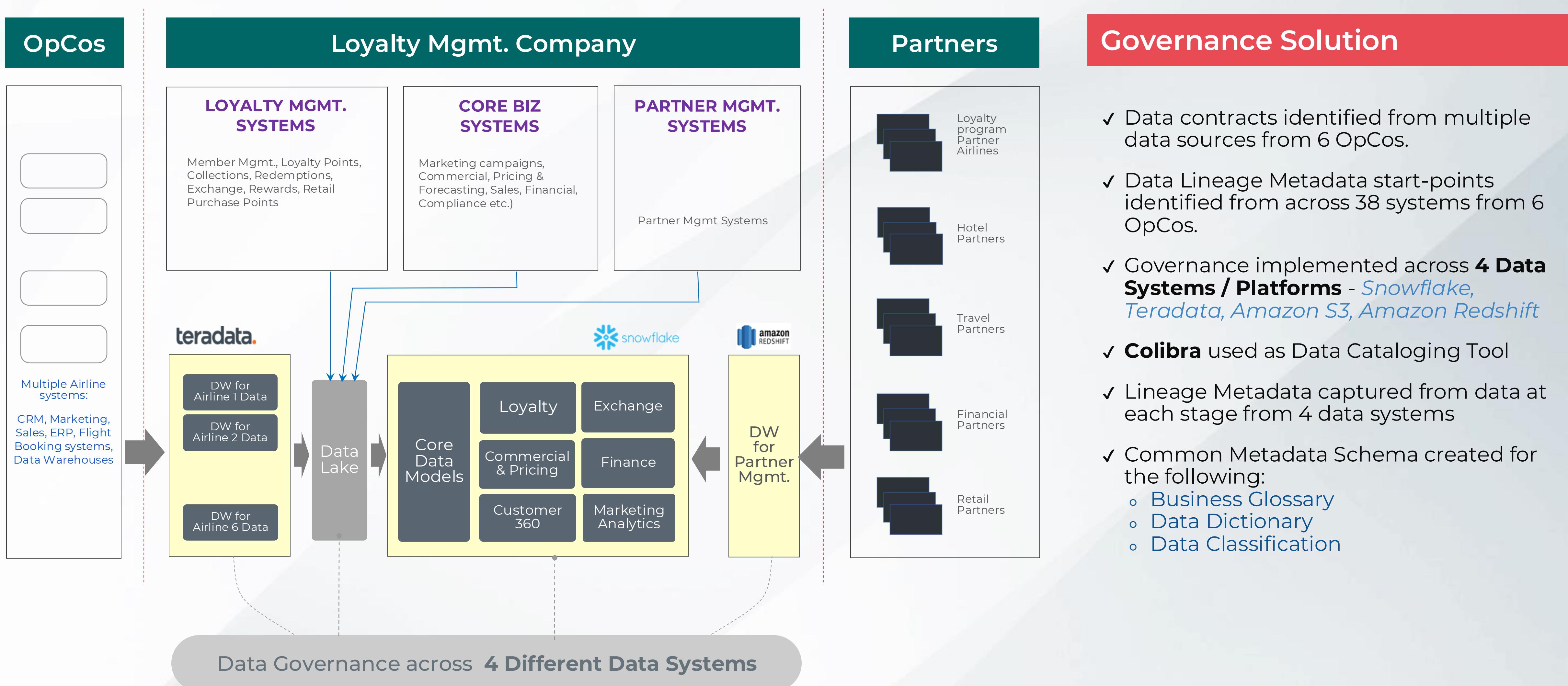
4+
Data Platforms

Partners

30+
Partners

30+
Partner Sources

Consolidation of multiple OpCos (2/2)



Data Governance at Scale: Health (1/2)



Company

Global Pop-Health Org serving 190+ member countries via 70 programs on 1500+ Health Indicators

Challenge

- Complexity in Data Ownership (since 70 programs owned the data)
- True Origin of Data was hard to map for Data Lineage
- Governance spanned to policies and rules for 'Infrastructure' and not just Data
- Challenging to create Business Glossary & Business Metadata as origin of data was hard to identify

Data Consumption from multiple Countries via multiple modes

Very Rigid Data Warehouse setup needing significant manual operations

Significant delays in generating Analytics

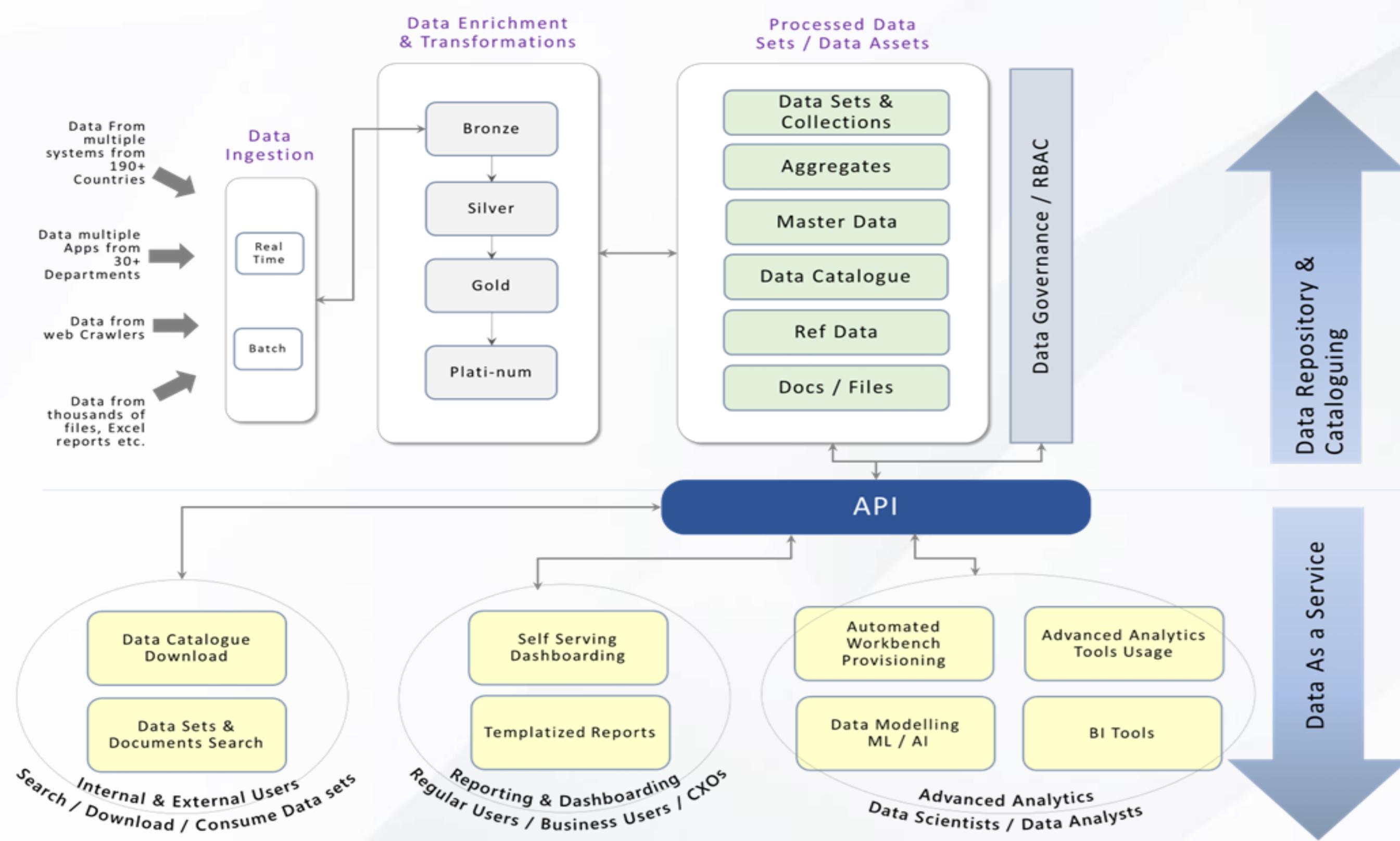


Zero Data Governance

No standardization of Data Modeling and Transformation

Managing PII and Sensitive Information was not handled

Data Governance at Scale: Health (2/2)



Governance Solution

- ✓ Governance achieved using a mix of Tools
- ✓ Source of the Data (country the department of origin) was manually updated
- ✓ **Azure Purview** was used to:
 - Catalog all data along all stages
 - Data Lineage
 - Business Glossary (manually updated)
- ✓ Azure Active Directory was used for Identification and Management and RBAC
- ✓ Azure Catalog Tags were used to Tag Azure resources (infrastructure governance)
- ✓ Azure DW was engineered to handle Security of Sensitive (e.g., PII) data



Azure
Data
Factory



Azure
SQL



Azure
AD



Azure
AD



Azure
AD

Governance for value and excellence (1/2)

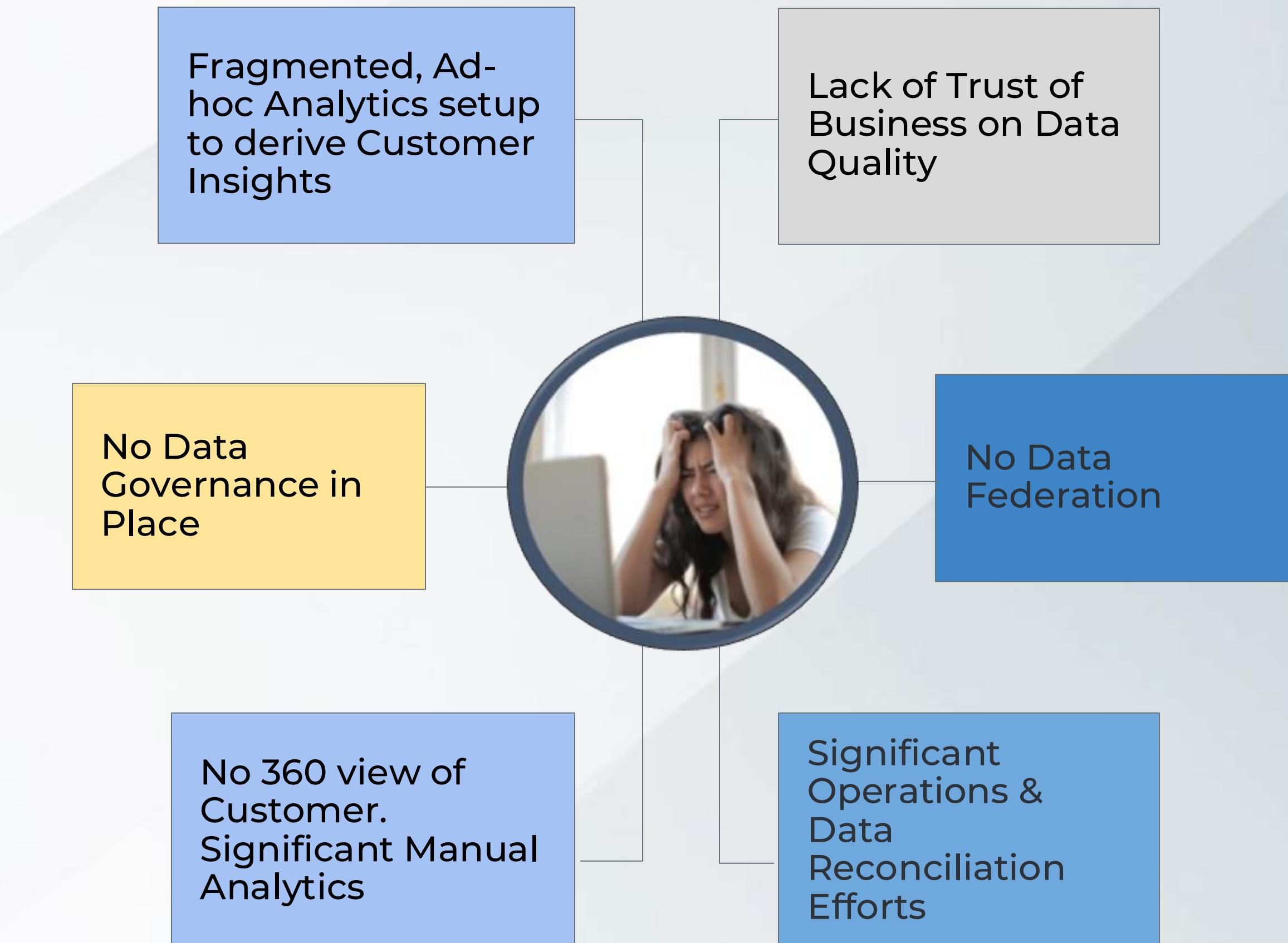


Company

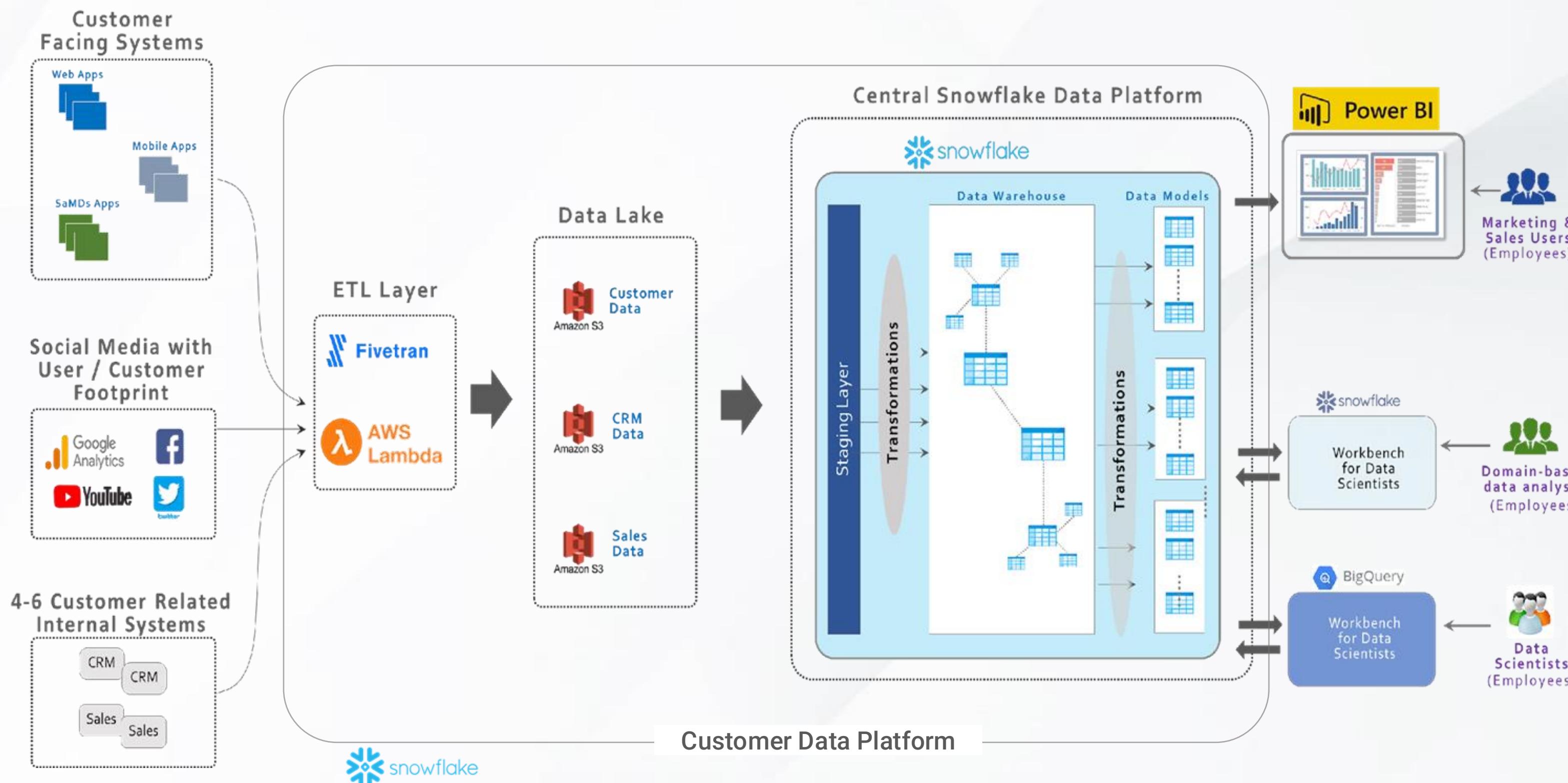
Global Medical Products and Technology leader in Advanced Wound Management and Ostomy Care, with 10,000+ workforce, 8 Manufacturing sites across the globe and operating in 100+ countries.

Challenge

- CDP (Customer Data Platform) without any data governance in place
- Insights from Customer (Patients, Doctors and Nurses) data spread across 100+ Customer-facing application (Web Portals, Mobile Apps, multiple CRM and Sales systems from multiple global locations).
- Lack of Data Lineage, Governance, Quality, Discoverability was becoming a major challenge from Regulatory & Compliance perspective.
- Huge amount of time and cost were spent in data collation and reconciliation



Governance for value and excellence (2/2)



Governance Solution

- ✓ Colibra used as Data Cataloguing Tool
- ✓ Following Governance Aspects were implemented:
 - Data Catalog
 - Data Quality
 - Data Lineage
 - Data Collaboration
 - Data Privacy





MDM Case Studies

1. GenAI based MDM Transformation for Distribution Business



1. Centralized

- **AI-Driven Ingestion & Unification:** Consolidated product data from multiple sources with automated normalization.
- **Generate & Enhance Attributes:** Identified missing attributes and enriched metadata using LLMs.
- **Graph/MDM Repository Setup:** Created a unified repository for seamless data access and semantic analysis.

2. Enriched

- **Semantic & Contextual Enrichment:** Mapped relationships and enhanced data discoverability with AI.
- **Natural Language Query:** Enabled intuitive search and auto-generated insights for teams.
- **AI Dashboards & Insights:** Delivered self-service analytics with AI-driven summaries and visualizations.

3. Visualized

- **Interactive GigaMap:** Illustrated product relationships and dependencies in dynamic maps.
- **AI-Guided Link Detection:** Uncovered connections like complementary products and supplier overlaps.
- **Conversational Exploration:** Enabled chat-based navigation for instant data access.

4. Standardized

- **AI-Driven Taxonomy & Schema:** Automated the creation and refinement of taxonomies and classifications.
- **Automated Attribute Validation:** Verified consistency and resolved conflicts with AI.
- **Lineage & Glossary Maintenance:** Tracked data transformations and ensured consistent terminology.

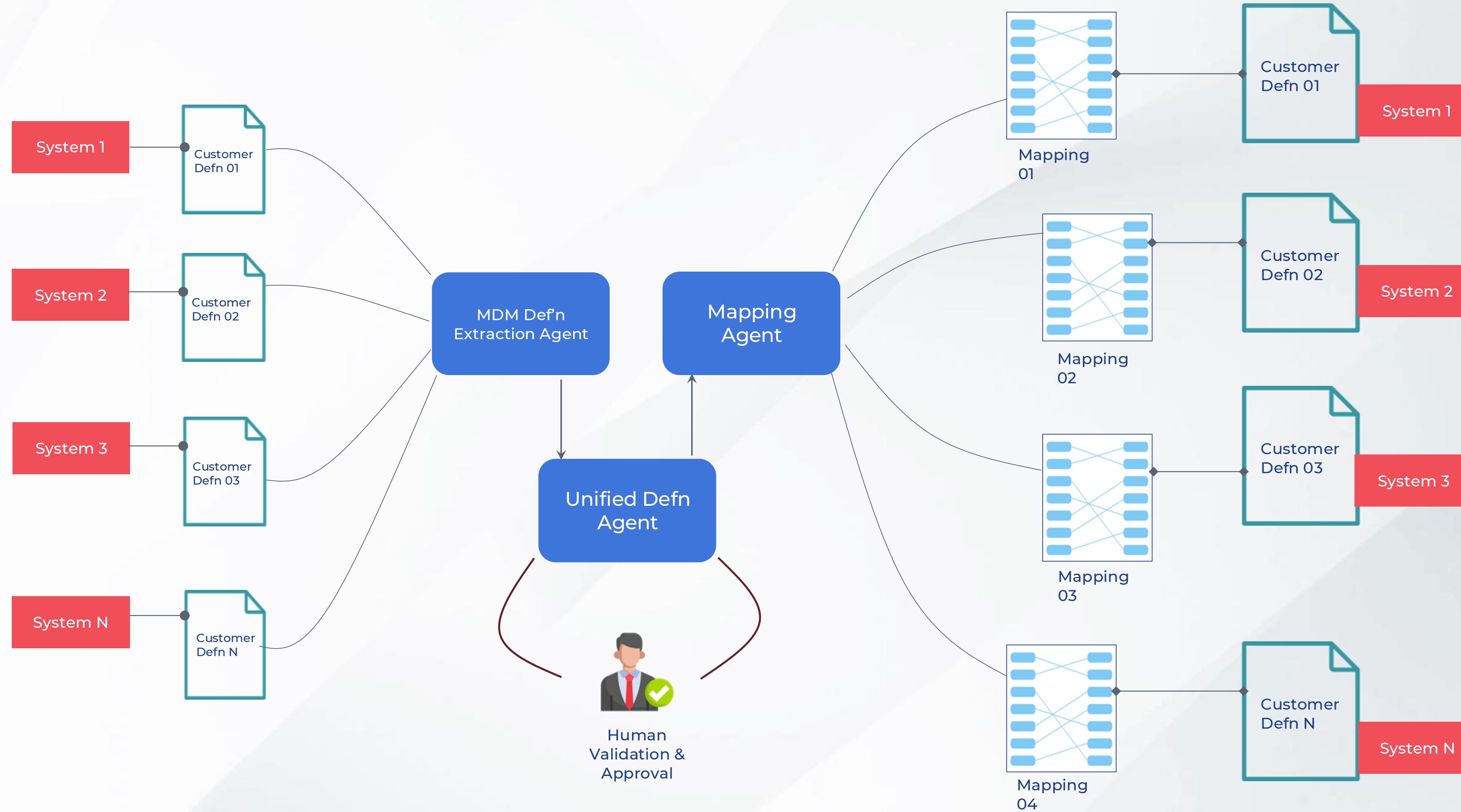
5. Operationalized

- **AI-Integrated Data Flows:** Embedded real-time enrichment into operational systems.
- **Automated Content Creation:** Generated product descriptions and marketing copy at scale.
- **Continuous Model Learning:** Adapted AI models to evolving data and product structures.

6. Governed

- **AI Recommendations for Ownership:** Suggested roles and responsibilities for data governance.
- **Policy & Compliance Enforcement:** Automated compliance checks and anomaly detection.
- **Data Quality Monitoring:** Delivered real-time alerts for data accuracy and consistency.

1. GenAI based MDM Transformation for Distribution Business



2. MDM Transformation for Medical Devices Company



Company Background

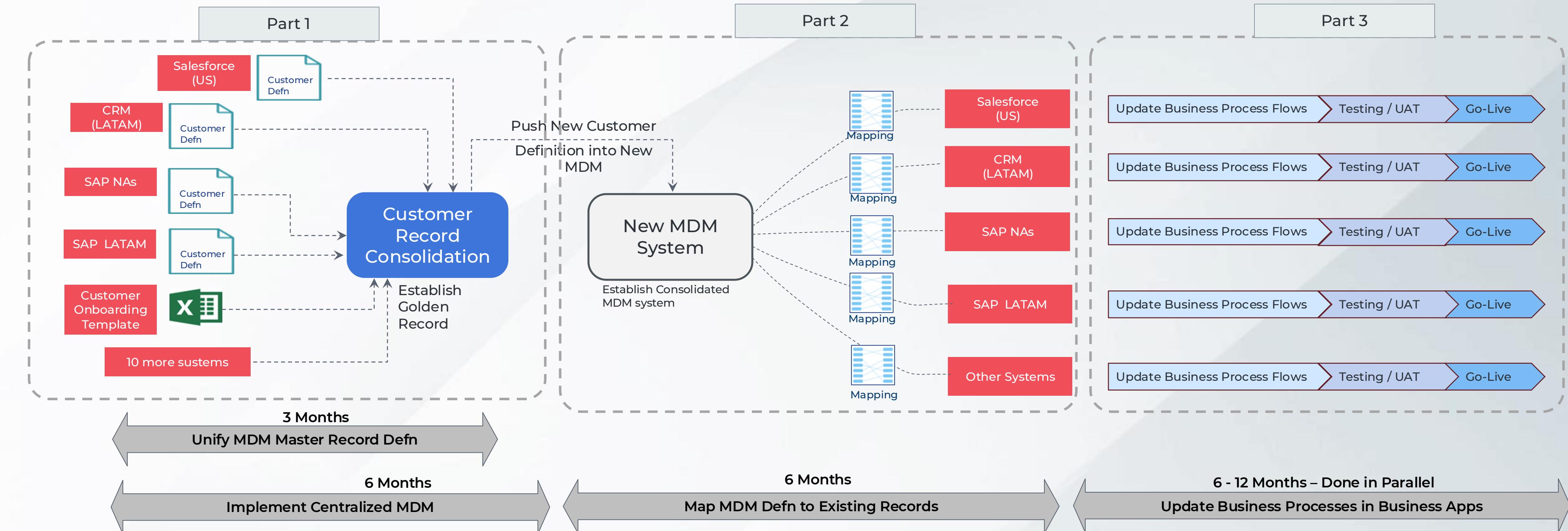
Global Life Sciences and Medical Devices Company with 3 main product lines, 10s of Thousands of SKUs and 15 Global locations (including 9 Manufacturing locations)

Challenge

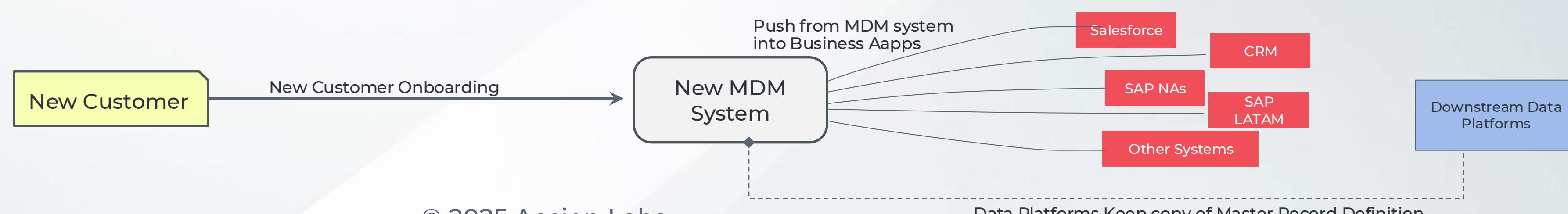
14+ ERPs, 6 CRM and 5 PLM systems produced significant variants of Master entities leading to tremendous loss of efficiency and reconciliation at various levels (processes, back-office operations and Data Analytics)

Solution Approach

- 2 Yr Program to Fix Customer and Product Master for Americas Geography
- Impact across Product line that generated 55% of total revenue
- Inventory Mgmt. Cost reduced by 50%
- MDM Transformation program carried in parallel to ERP transformation program



Post Go-Live



3. Product Master Transformation for CPG Industry



Problem Statement :

The organization seeks to establish a unified, authoritative 360-degree view of product data that currently exists in fragmented form across three distinct business tiers: manufacturers, distributors, and retailers.

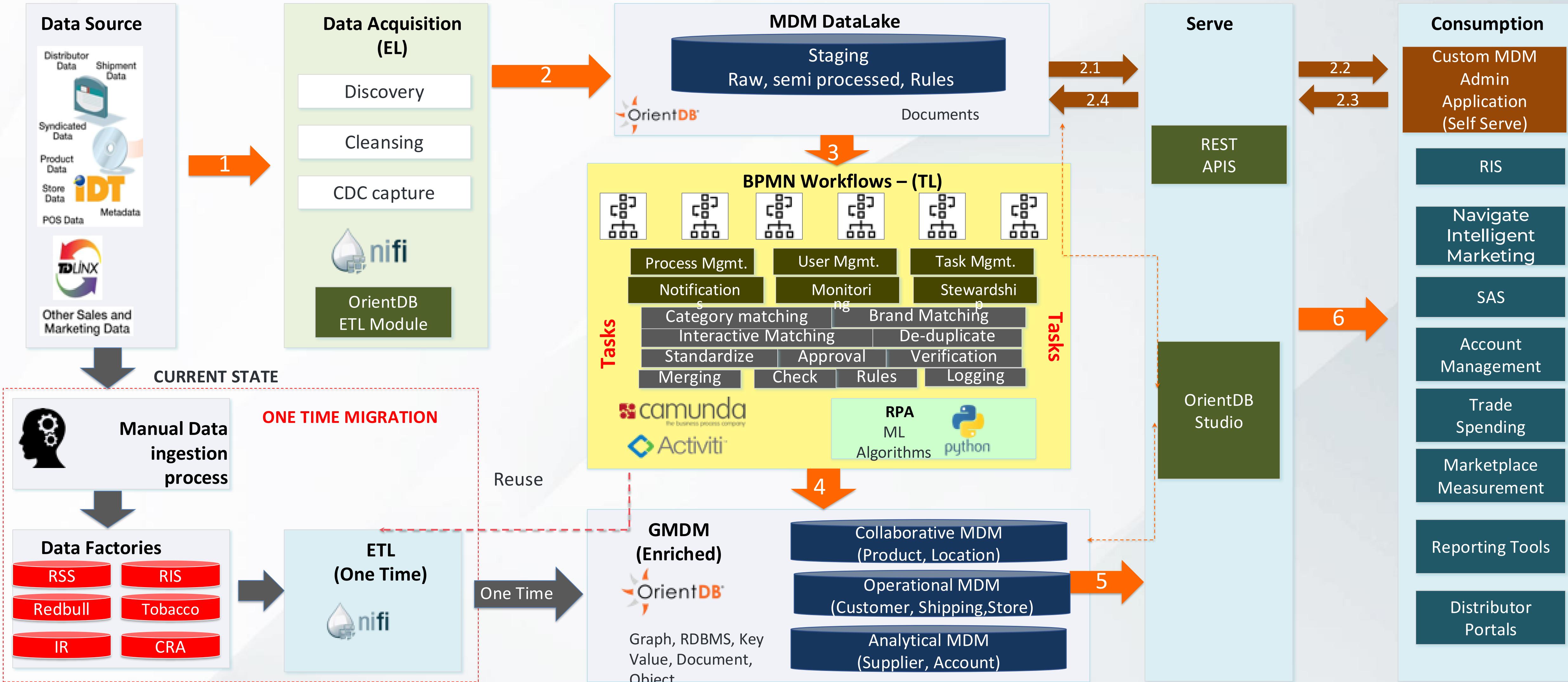
Primary Data Quality Considerations:

- Manufacturer's data serves as the highest quality source, providing the foundation for product attributes and specifications
- Distributor data presents moderate quality concerns with inconsistent formats and incomplete attribute mapping
- Point of Sale (POS) retail data exhibits significant quality issues, including variant product descriptions, incomplete specifications, and inconsistent categorization

The objective is to create a golden record for each product that:

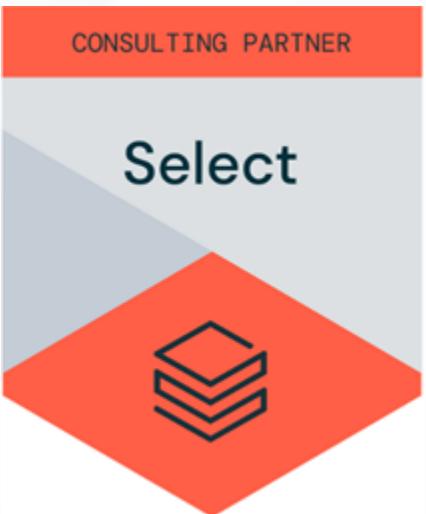
- Leverages the superior quality of manufacturer data as the authoritative source for core product attributes
- Enriches this foundation with valuable business context from distributor channels, including pricing tiers, regional availability, and channel-specific product variants
- Incorporates critical market performance indicators from retail POS data while implementing robust data quality measures to standardize and cleanse this information

3. Product Master Transformation for CPG Industry



Databricks

Our Databricks Credentials



65

Technical Trained

15

Technical
Certified

20

Clients

30

Lake Houses
Built

5

Gen AI
Solutions



OUR SERVICES

DIGITAL PRODUCTS



- Microservices and Cloud-native Architecture and Development
- Product Engineering and Application Development
- Kubernetes and Containerization Architecture and Implementation
- API, iPaaS & Digital Integration
- Application Re-engineering, Data Migration, Deployment QA and Testing

CLOUD ADOPTION & DEVOPS



- Cloud Migration Strategy and Planning
- Cloud Architecture, Governance, Security and Architecture
- DevOps, CI/CD Design and Implementation
- Infrastructure as Code and Deployment Automation
- Datacenter, Server and Database Migration



DATA & ANALYTICS

- Data Platform Modernization Strategy & Design
- EDW to Data Lake/Lakehouse Migration
- AI/ML, Predictive Analytics Design and Development
- Data Engineering, Data Processing, Data Cleansing/Conversion Development
- MDM/ Data Governance Roadmap
- Dashboard, Power BI and Reporting



AUTOMATION & OPERATIONS

- RPA and Process Automation
- SRE & Cloud Operations
- Salesforce & Service Now implementation
- QA & Test Automation
- SLA-backed Managed Services

CLOUD AND INNOVATION ENGINEERING SERVICES



Our Azure Data & AI Offerings



CLOUD-SCALE ANALYTICS

Power Business Decisions with Azure Data Platform

- Data Lake & Data Warehouse
- Lake House
- BI Reporting, Dashboard and Analytics
- Streaming data Analytics

- Azure Databricks
- Microsoft Fabric
- Azure Synapse Analytics
- Azure SQL MI, CosmosDB

UNIFIED DATA GOVERNANCE

Govern, protect, & manage your entire data estate

- Data Catalog and Classification
- Data Quality
- Data Lineage
- Master Data Management

- Microsoft Purview
- Profisee

AI & ML OPS

Turbocharge your GenAI and AI/ML initiatives

- Risk Analytics
- Cognitive Services
- Document, Audio & Video Processing and analytics
- NLP & Sentiment analytics

- Azure Open AI
- Azure Cognitive Services
- Azure ML Service

EDW / DWH MODERNIZATION

Modernize your Data Estate with Azure Data Platform

- Re-engineer EDW Workloads
- Enterprise DBs to Cloud DWH
- On-prem to Azure Migration
- Spark Data Engineering
- SQL Server, Oracle, Hadoop, MySQL, PostgreSQL

- Azure Databricks
- Microsoft Fabric
- Azure Synapse Analytics
- Azure SQL MI

Data ISV's Partnerships



How Accion Labs works with Databricks



Greenfield Data Platform Build and Sustenance

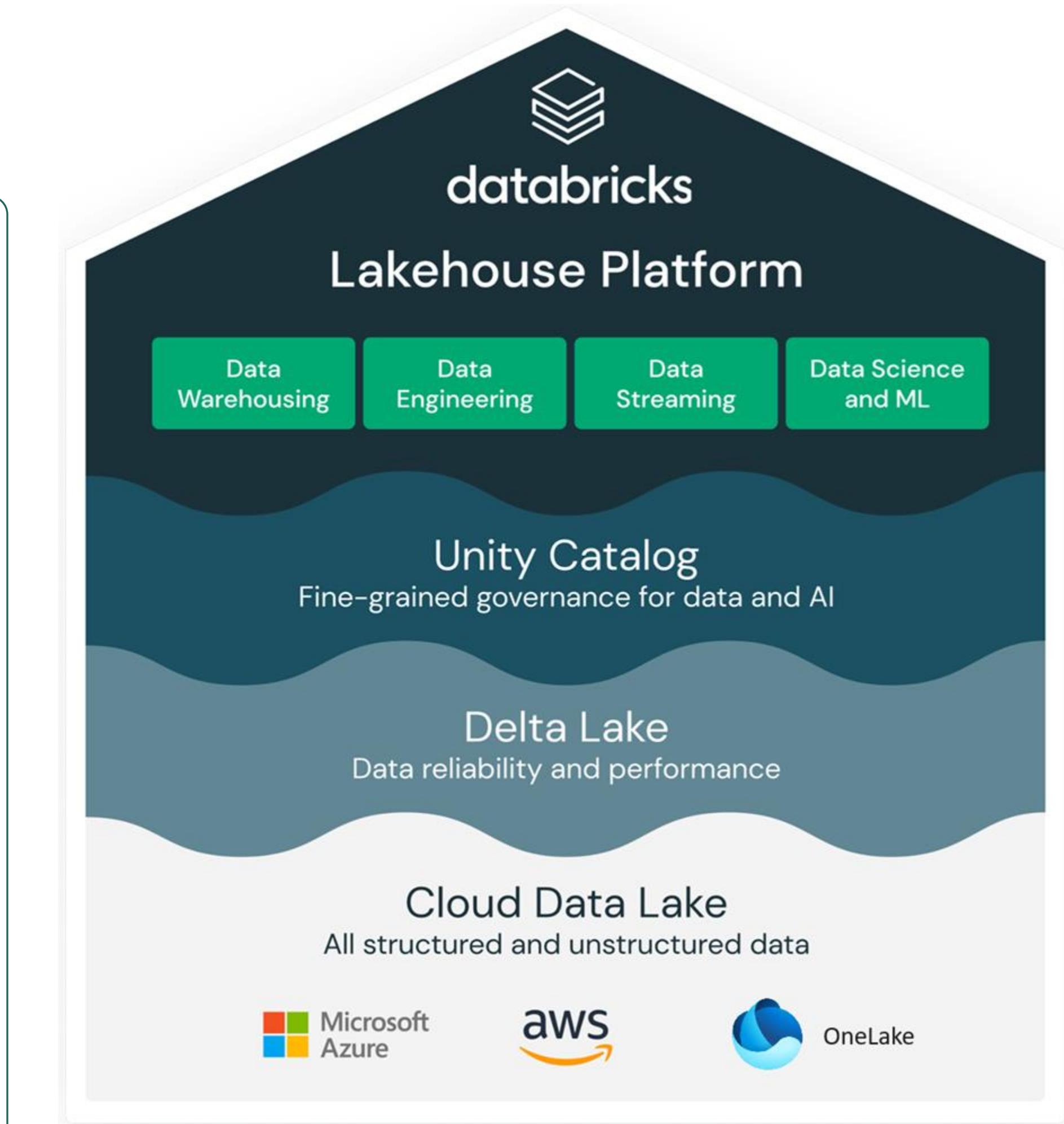
- Build a data platform from scratch
- Gather and Understand Requirements
- Build a roadmap
- Iterative Build and Sustenance of Layered Data Platform
- Build and Deliver Applications
- Enable Data Quality and Governance

Data Lake/DW Modernization and Sustenance

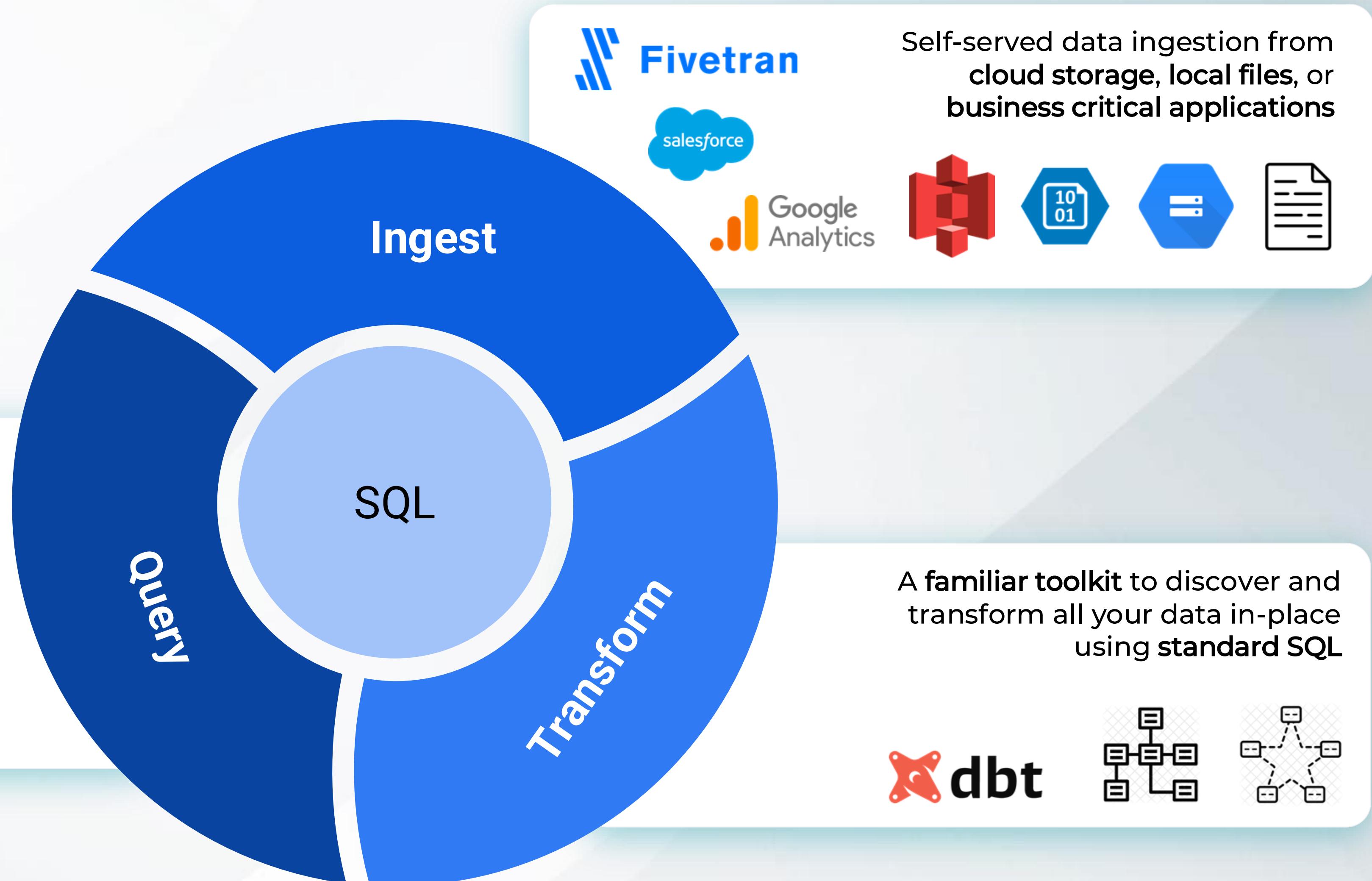
- Gather and Understand current and upcoming needs
- Perform a fit gap analysis with existing ecosystem
- Prioritize and build a roadmap
- Iterative Accelerated Build using accelerators wherever possible to convert
- Downstream compatibility enabled iterations
- Build and Deliver Applications

Advanced Analytics Solution Build and Sustenance

- Curate and Build Datasets based on the applications
- Build the Analytics models
- Build the applications or integrate datasets back into the data platform
- Perform Sustenance of the applications



Ingestion with Databricks



Unity Catalog Implementation



Delegation of Management (admin isolation)

Each SDLC environment has its own admin

Workspace to catalog binding

The PRD workspace is fully isolated and has its own catalog. DEV workspace has access to DEV and STG catalog

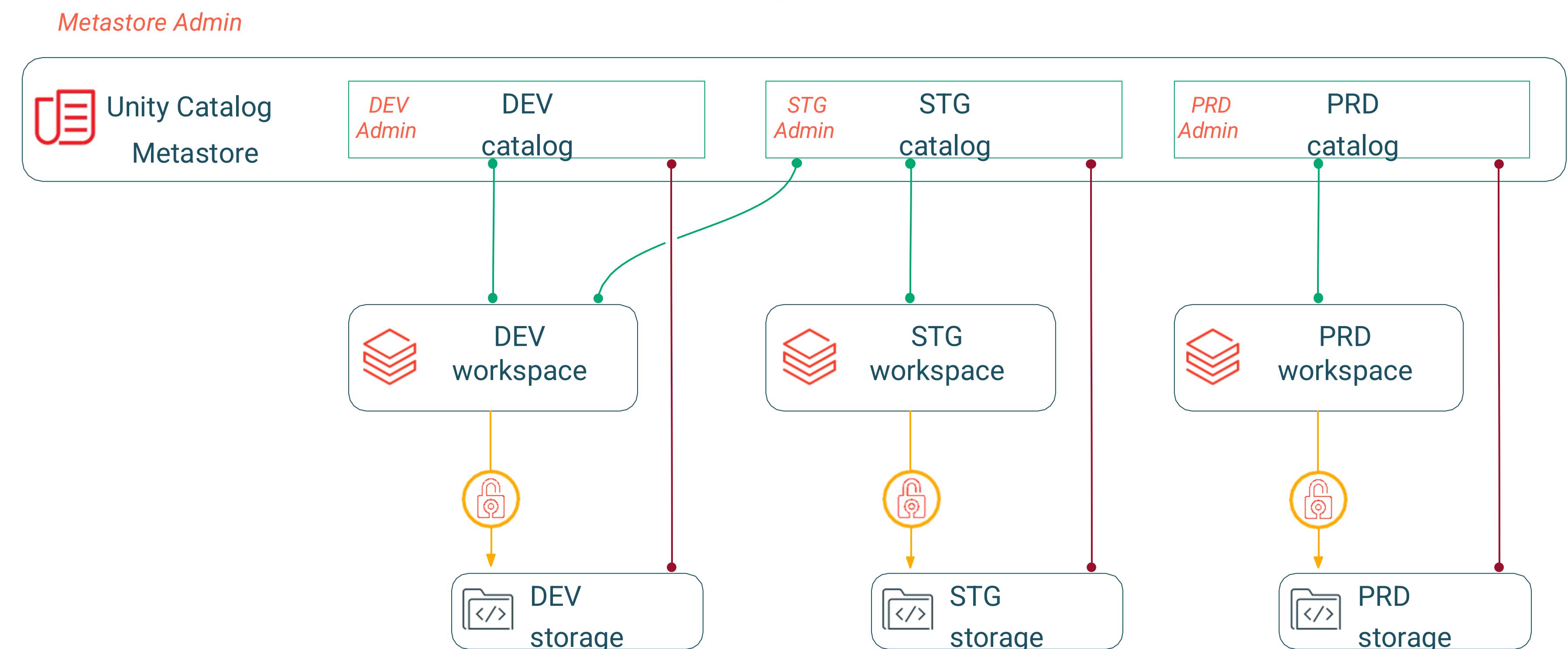
Storage isolation

Separates the storage locations on catalog level (typically sufficient)

UC Access Control

Users should only gain access to data/metadata based on agreed access rules

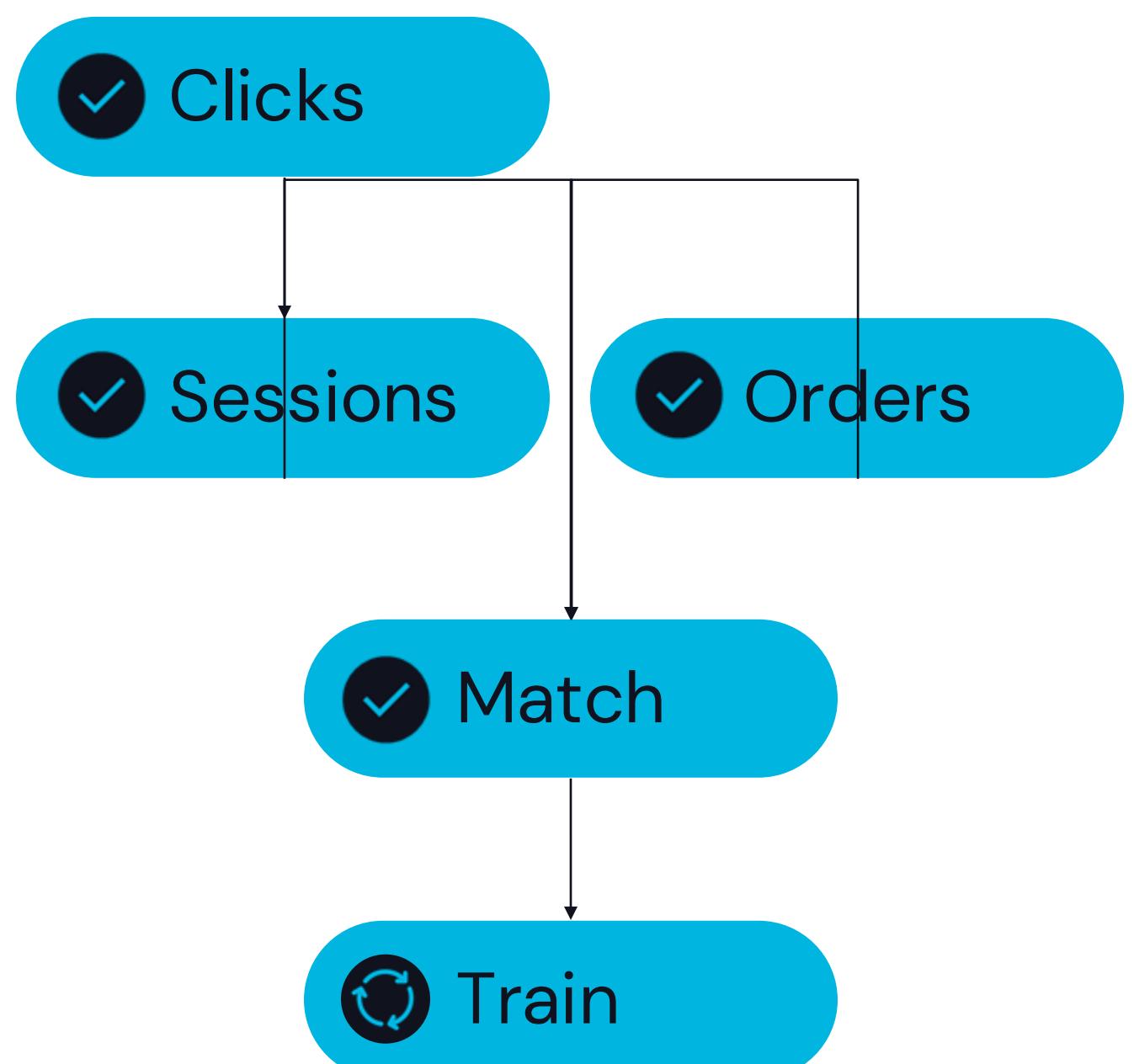
Cross-Workspace



Databricks Workflows

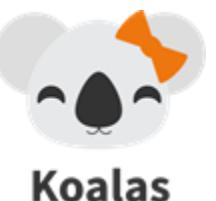


Orchestrate...



...any task...

Delta Live Tables



Auto Loader



mlflow™



non-Spark

dbt

...across any platform



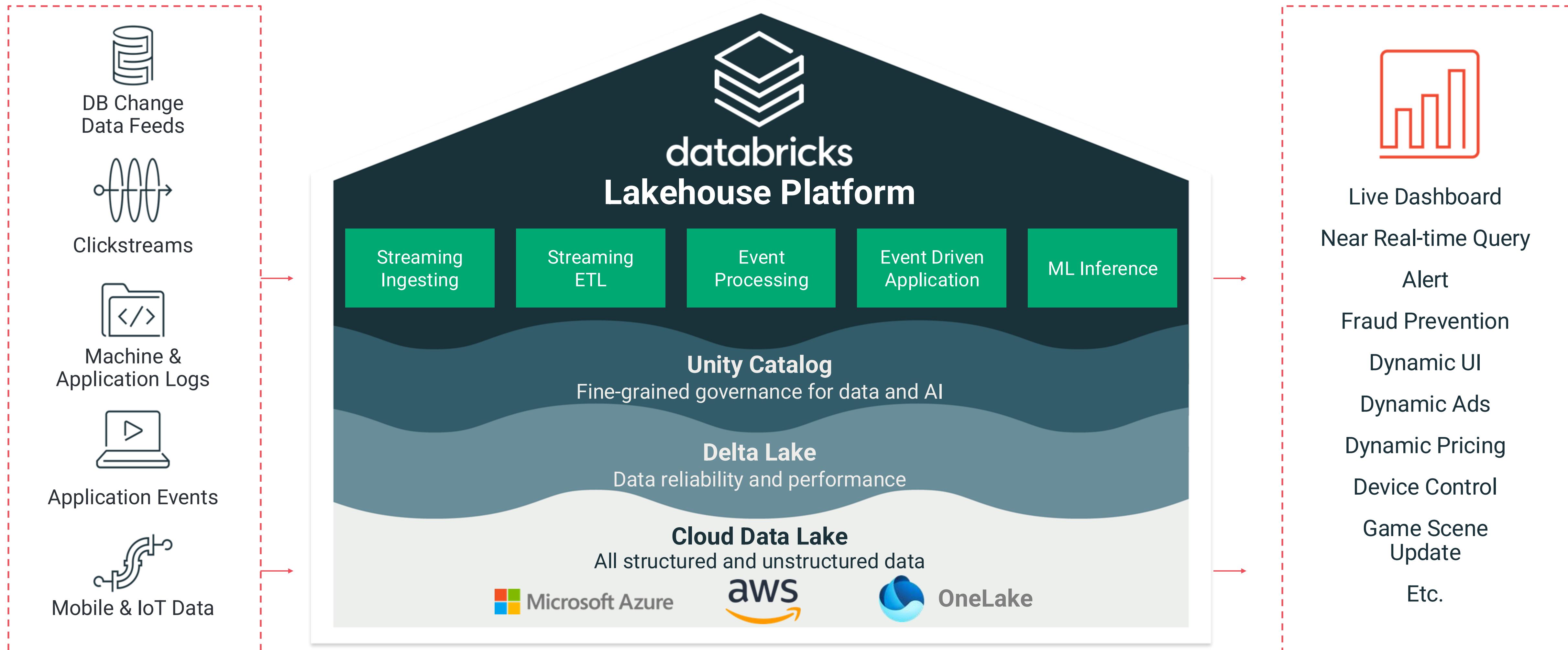
talend



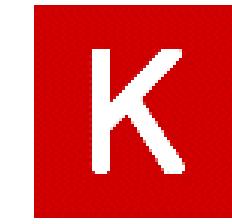
+more

+more

Streaming Applications (with Structured Streaming)



Building ML using Databricks



Model Building and Training



Model Tracking and Registry



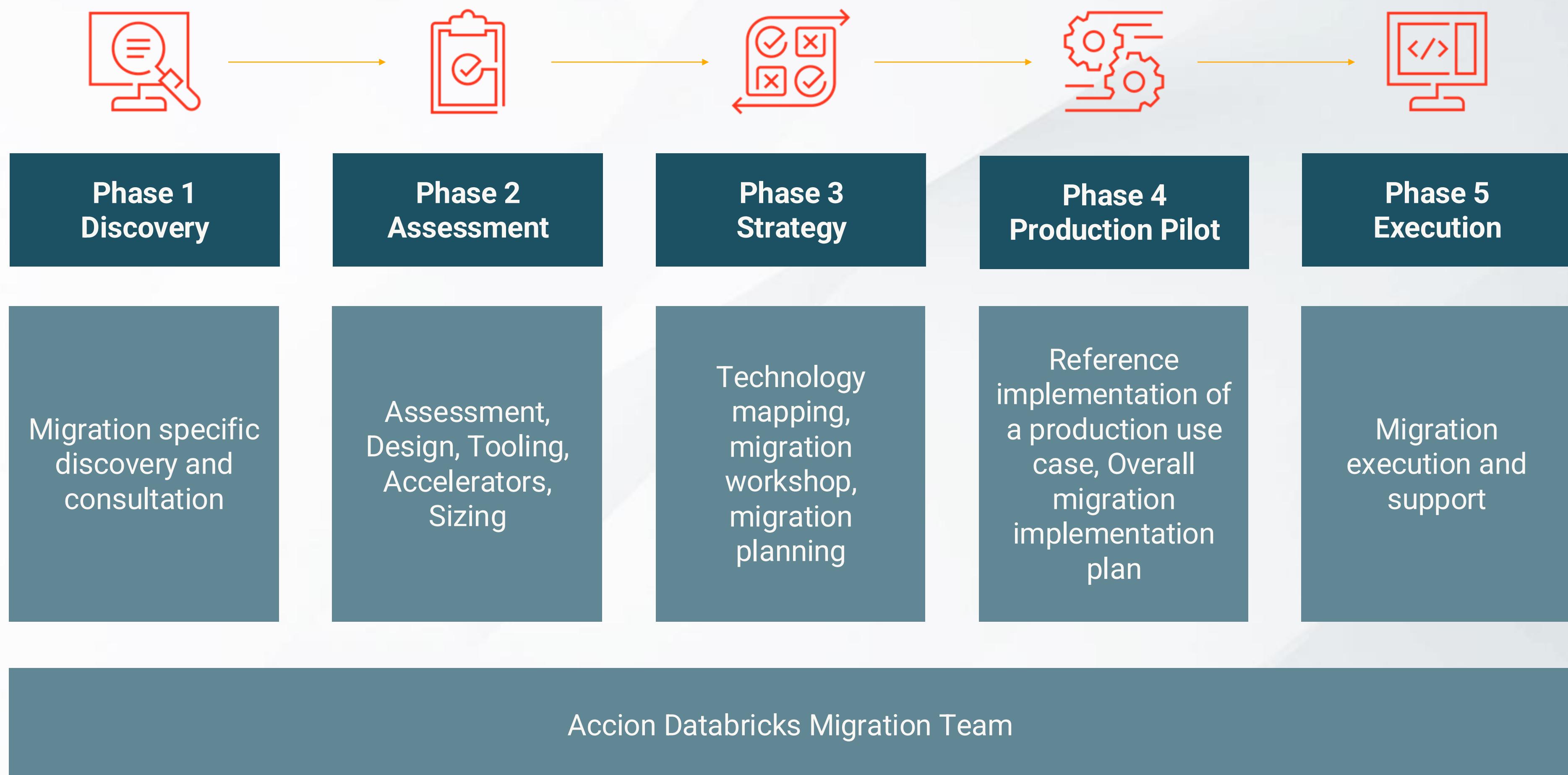
Runtime and Libraries



Automation and Governance

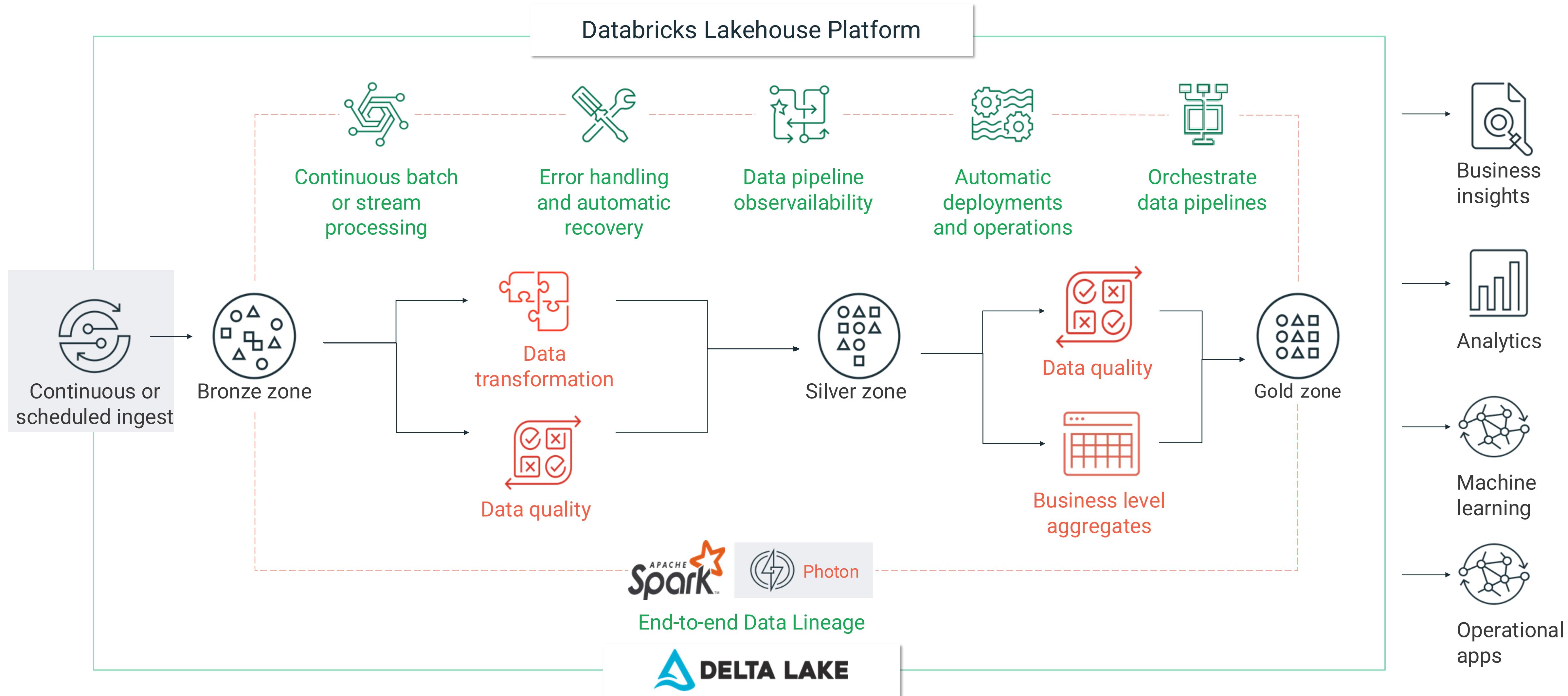
Modernization Methodology and Approach

Accion Modernization Methodology



Accion Databricks Migration Team

Modern data engineering in Databricks Lakehouse

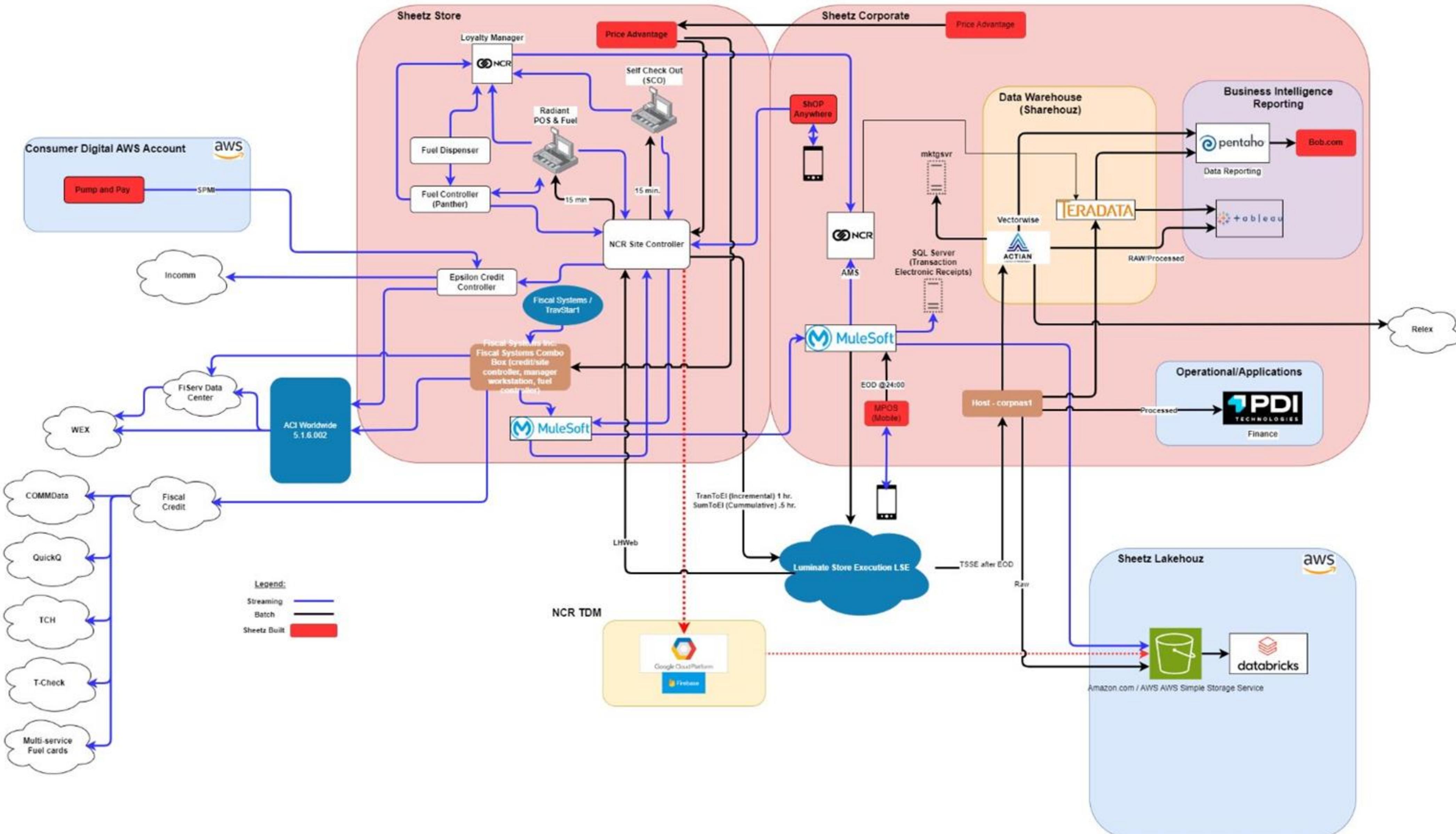


Select Databricks Case Studies

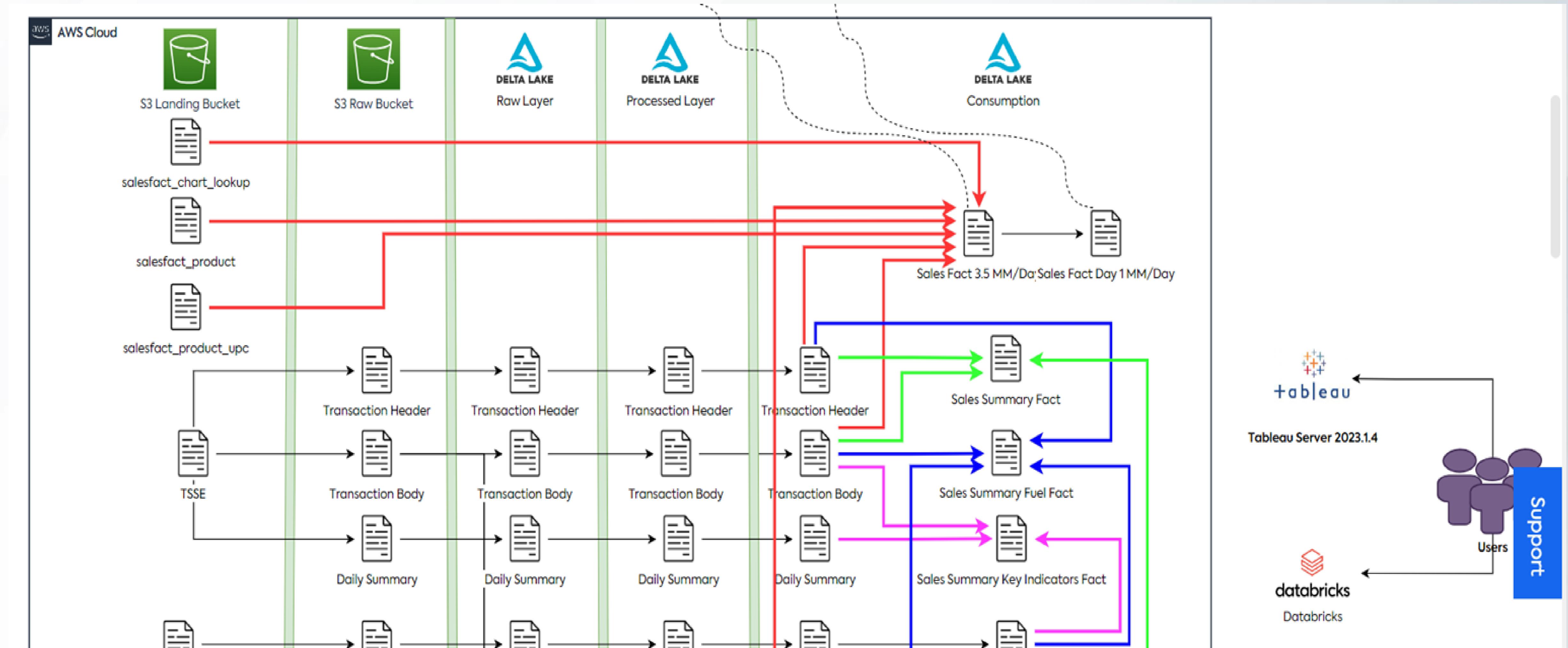
Sheetz – World before Databricks



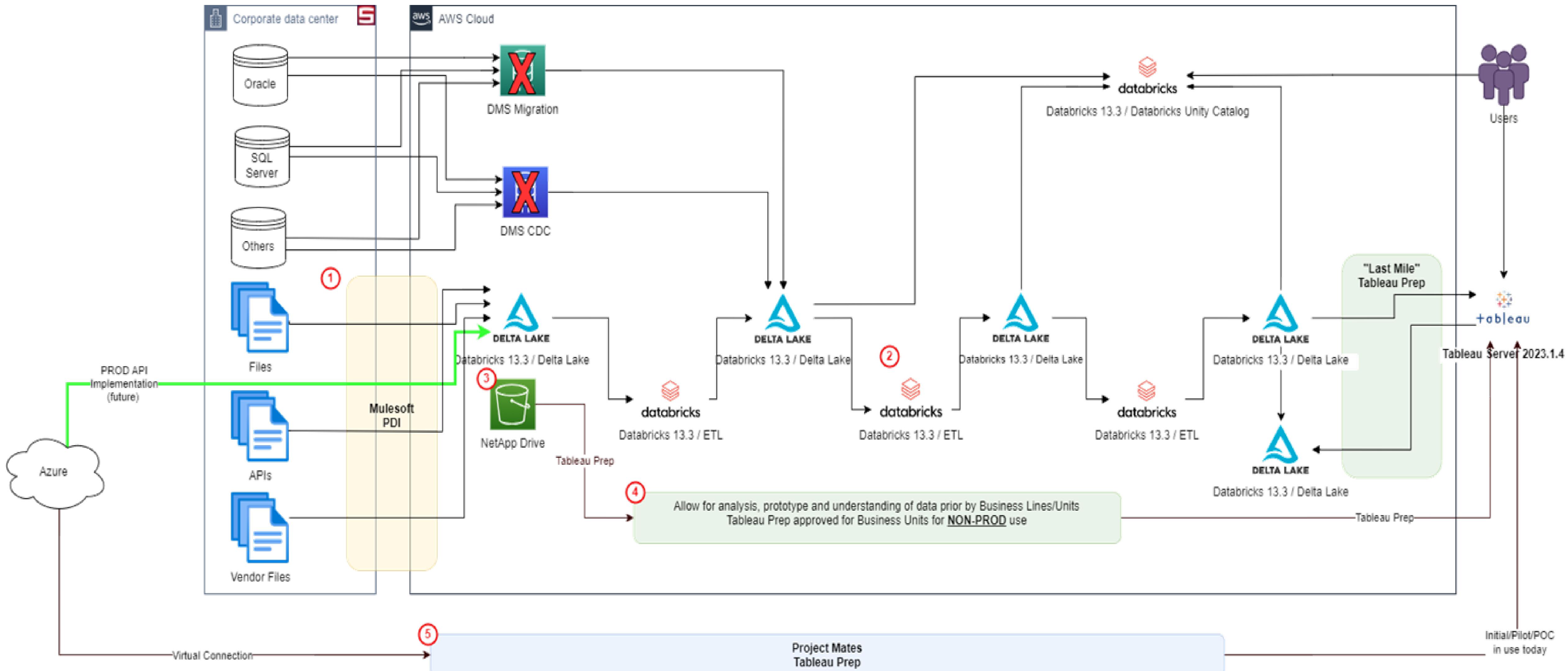
Note 1: Pump / Store Car Wash Purchase



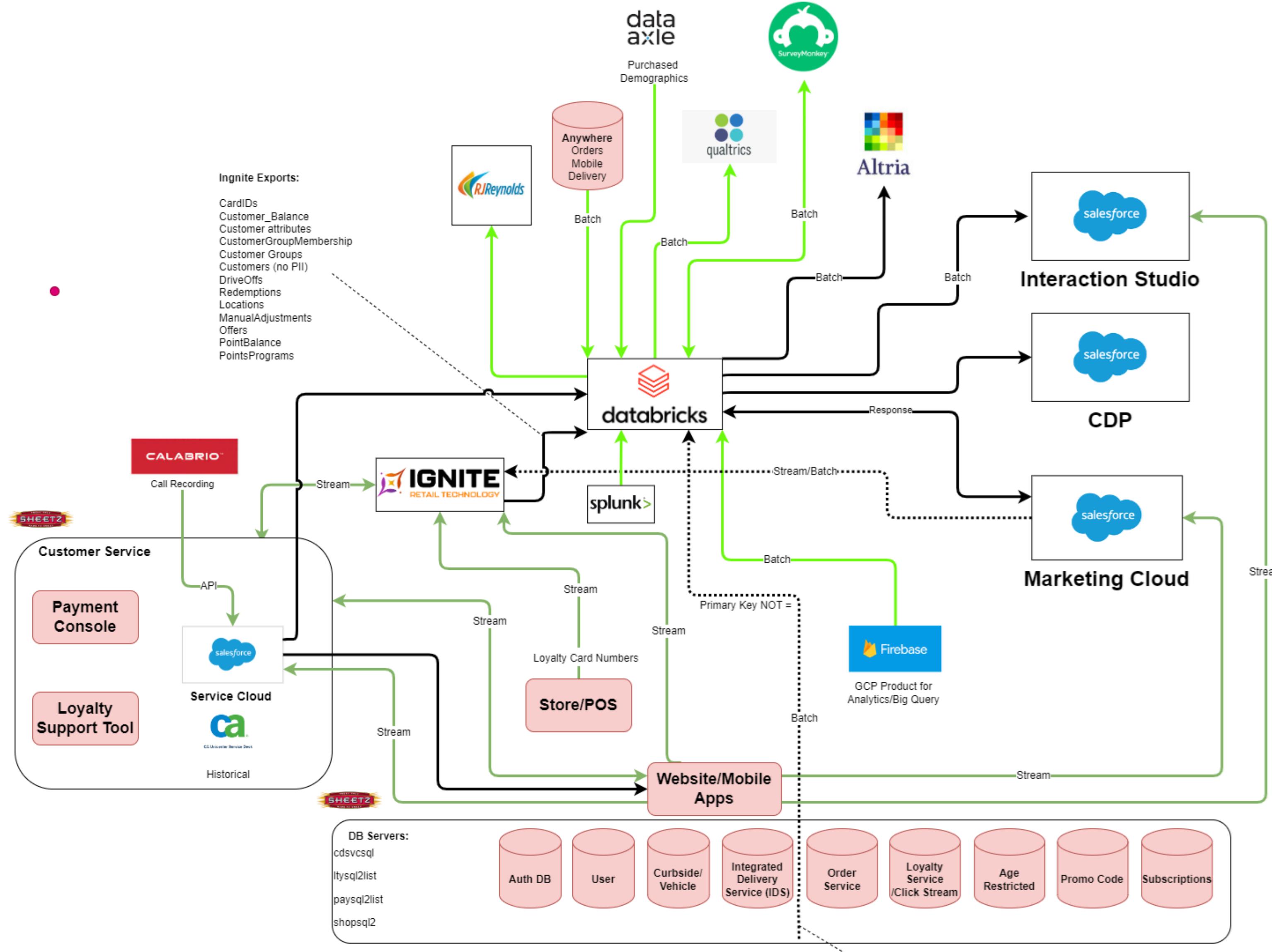
Sheetz – Adopting Databricks for POS Sales Data



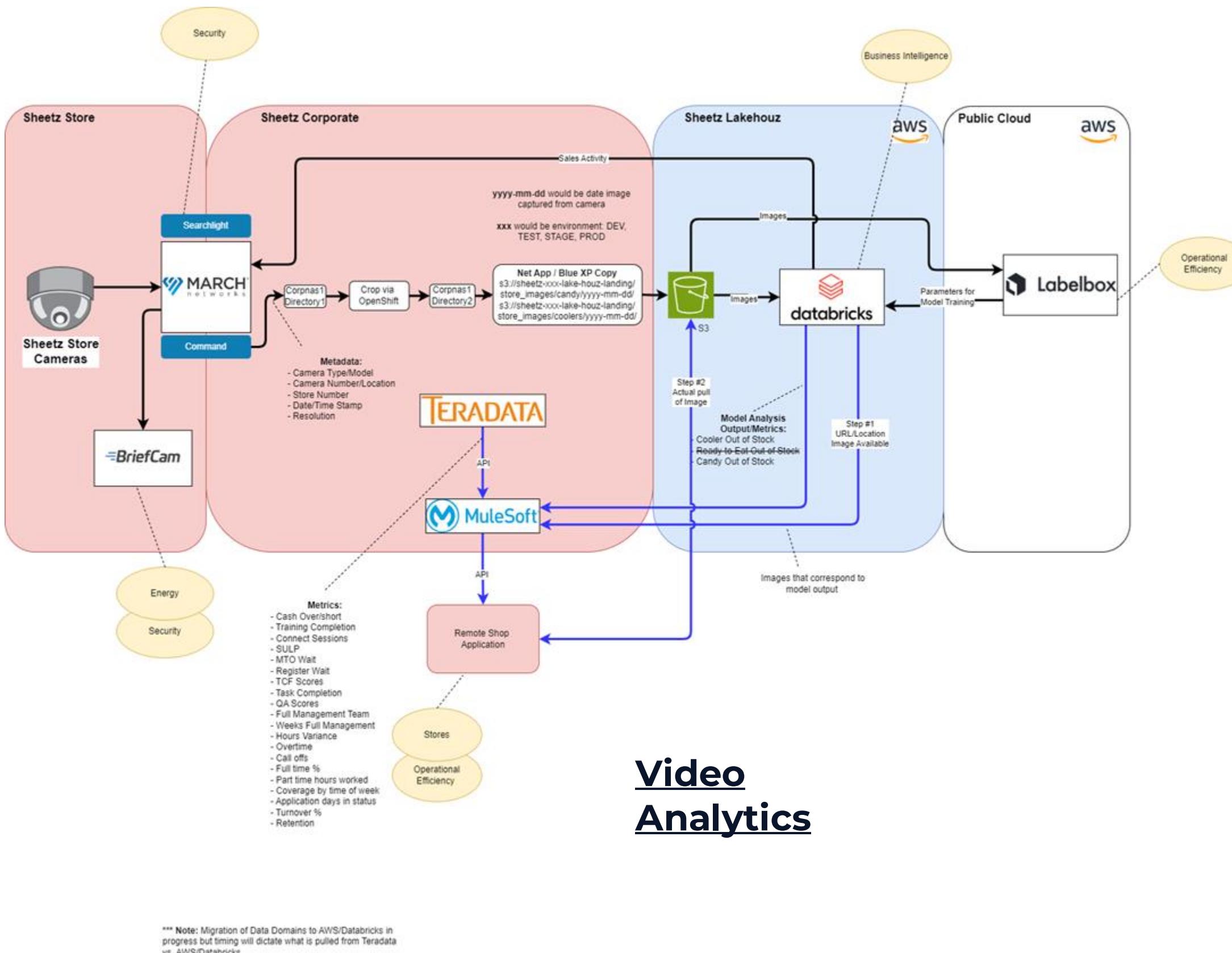
Sheetz – Embracing Databricks First



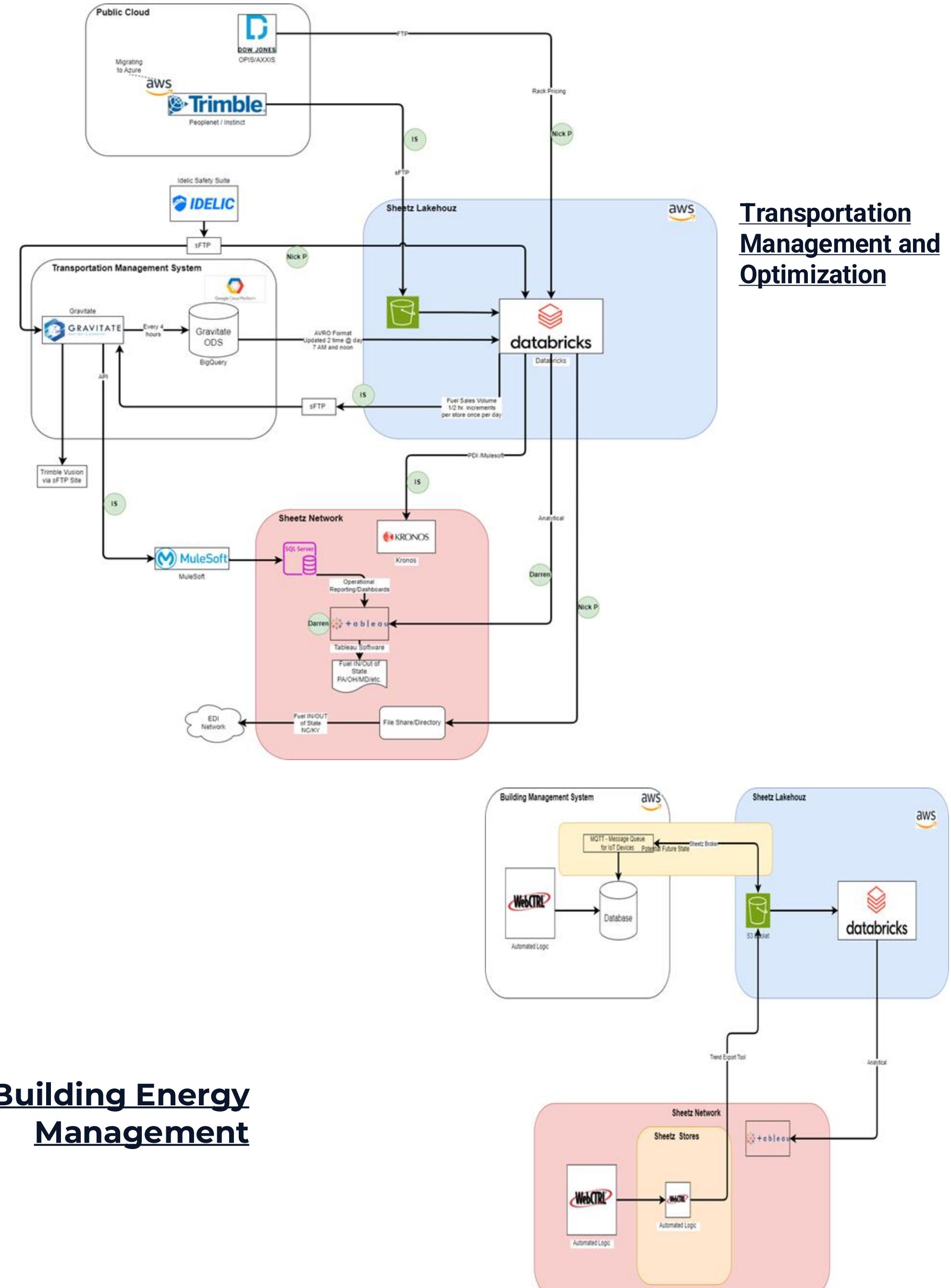
Sheetz - Modernized Enterprise View



Sheetz - Other Use Cases



Video Analytics



Building Energy Management

Driving Success using Databricks – Select Case Studies



Industrial Printers
Predictive Maintenance,
Sales forecasting .
Databricks Delta Lake, Delta Live Tables, Power BI



Financial Advisory and Consulting
Delinquency ML Model, Gen-AI using
Mosaic, Data Consolidation
Databricks Lakehouse, Mosaic, ML Models, Delta Tables, SAP, ServiceNow, Salesforce



Community Bank and Financial Services
Warehouse Migration and Modernization
Databricks Lakehouse, Unity Catalog, AWS Redshift, Google BigQuery, Looker, Power BI



Leadership Development
Course Completion and
Revenue Prediction
Databricks Delta Live Tables
Unity Catalog, Oracle ERP, D365 CRM, Marketo, DRIFT



Global Publication House
Business Warehouse with Self-service Reporting capabilities
Databricks Delta Live Tables, Databricks Lakehouse, SQL Server, Power BI, Marketo, Open Library



Educational Programs
Warehouse Migration
Databricks Lakehouse, SQL Server, Workflows



Compressor Manufacturing
SAP HANA to Databricks Migration
Databricks Lakehouse, Workflows, Delta Tables, Power BI, SAP HANA, Multiple SAP Environments



Healthcare ATS Platform
Candidate Prioritization, Candidate Profile Matching, Natural Query for searches
Databricks Lakehouse, Databricks SQL, Databricks Workflows, Unity Catalog , Bullhorn ATS, Azure OpenAI, Redis



Telecommunication Services
Bills generation, Complex Events Processing (Streaming data)
Databricks Lakehouse, Spark Streaming CEP, Databricks as ODS,

Streaming Case Studies

SOLUTION
IoT, Digital Platform

FOCUS AREAS
**Process Automation,
IT-OT, Operational
Efficiency, Accounting**

INDUSTRY
Oil & Gas, Logistics

Unique automated fuel delivery and management technology that uses advanced sensors to collect over 500,000 data points per day, providing real-time visibility of operations, guaranteed fuel supply and new levels of efficiency

\$1.4M

Net Operating Savings
in first 6 months

85%

Reduction in invoice times
Down to 2 days from 30 days

\$7.5M

Working Capital Improvements
in first 6 months



1. FAS Units with
IoT sensors



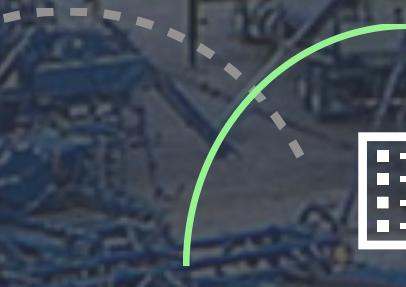
2. Automated fueling
with 28 hoses



3. Remote
Monitoring



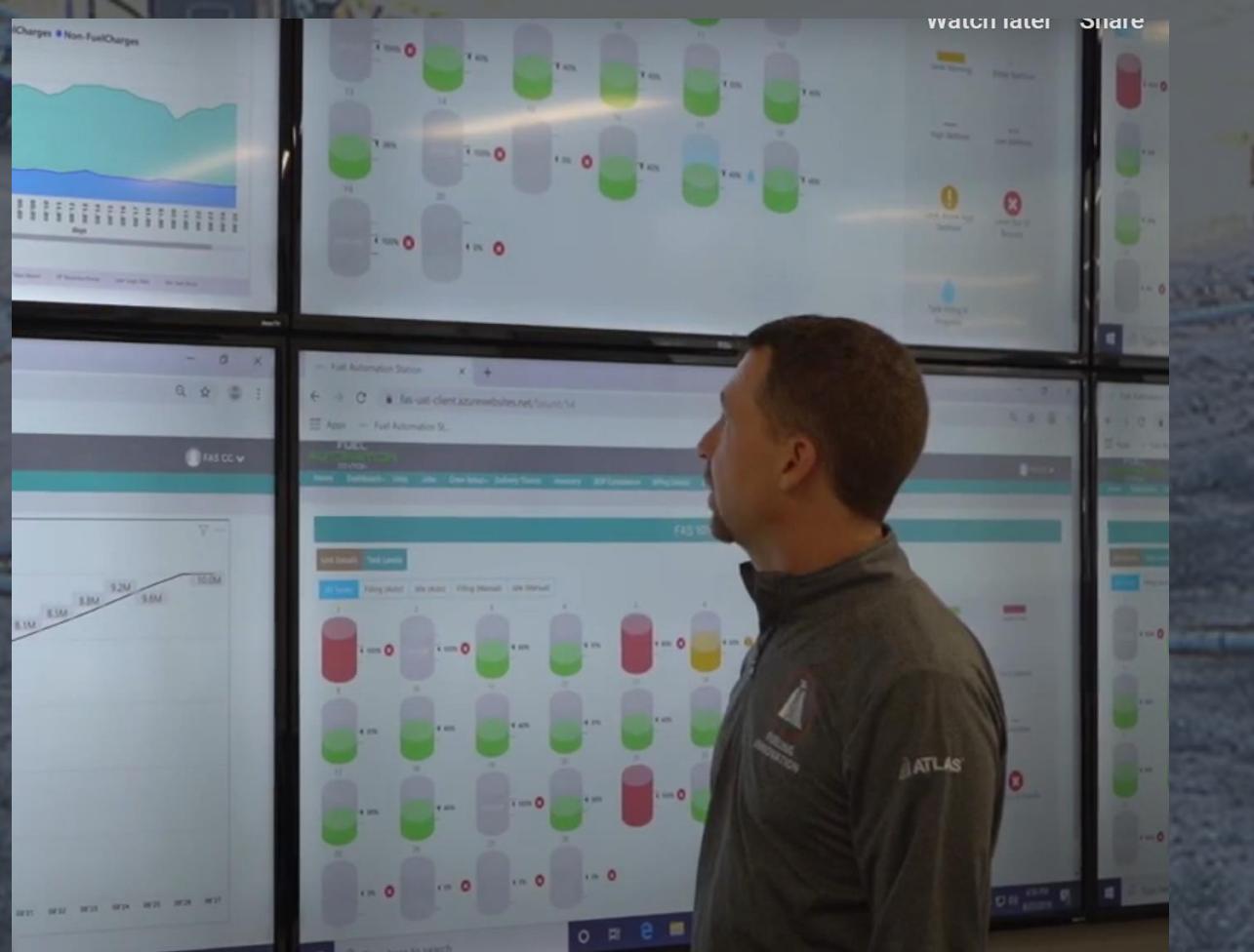
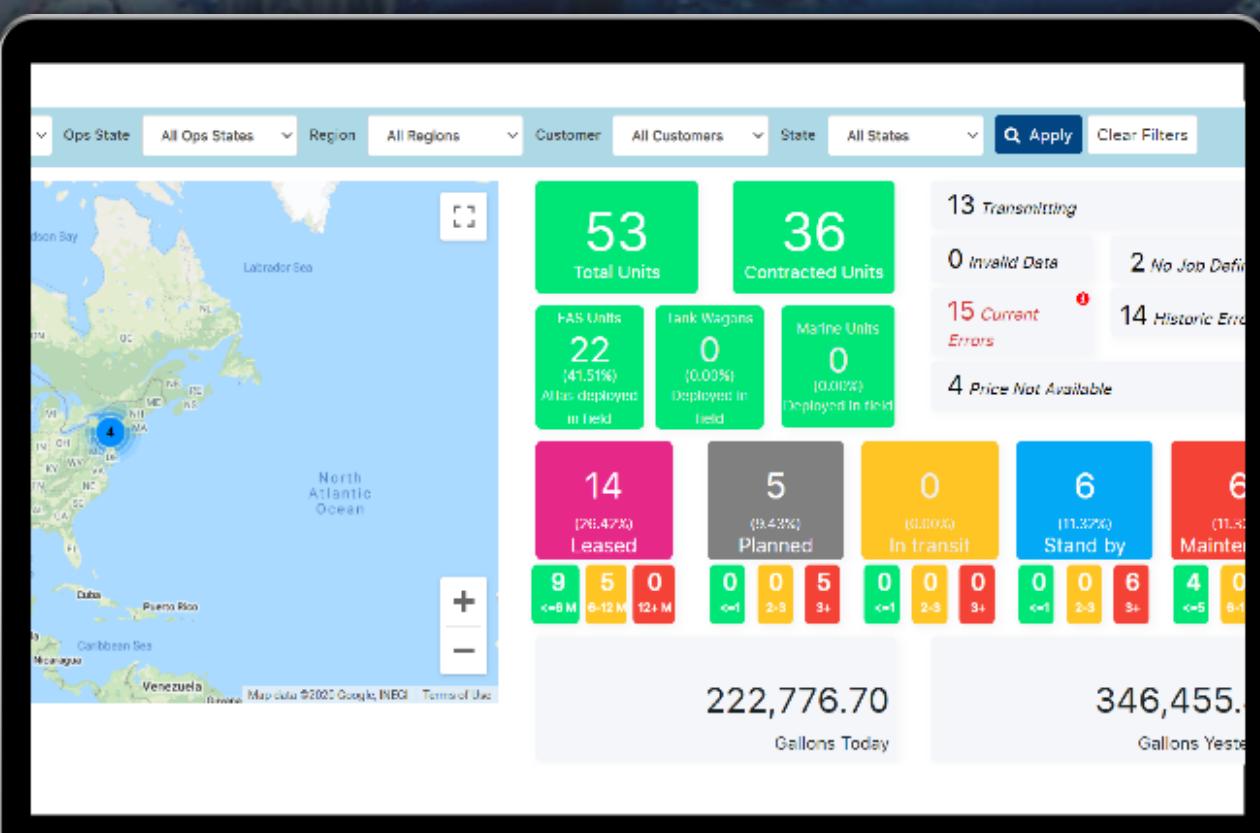
4. Real-time
tracking



5. Automated invoice
generation

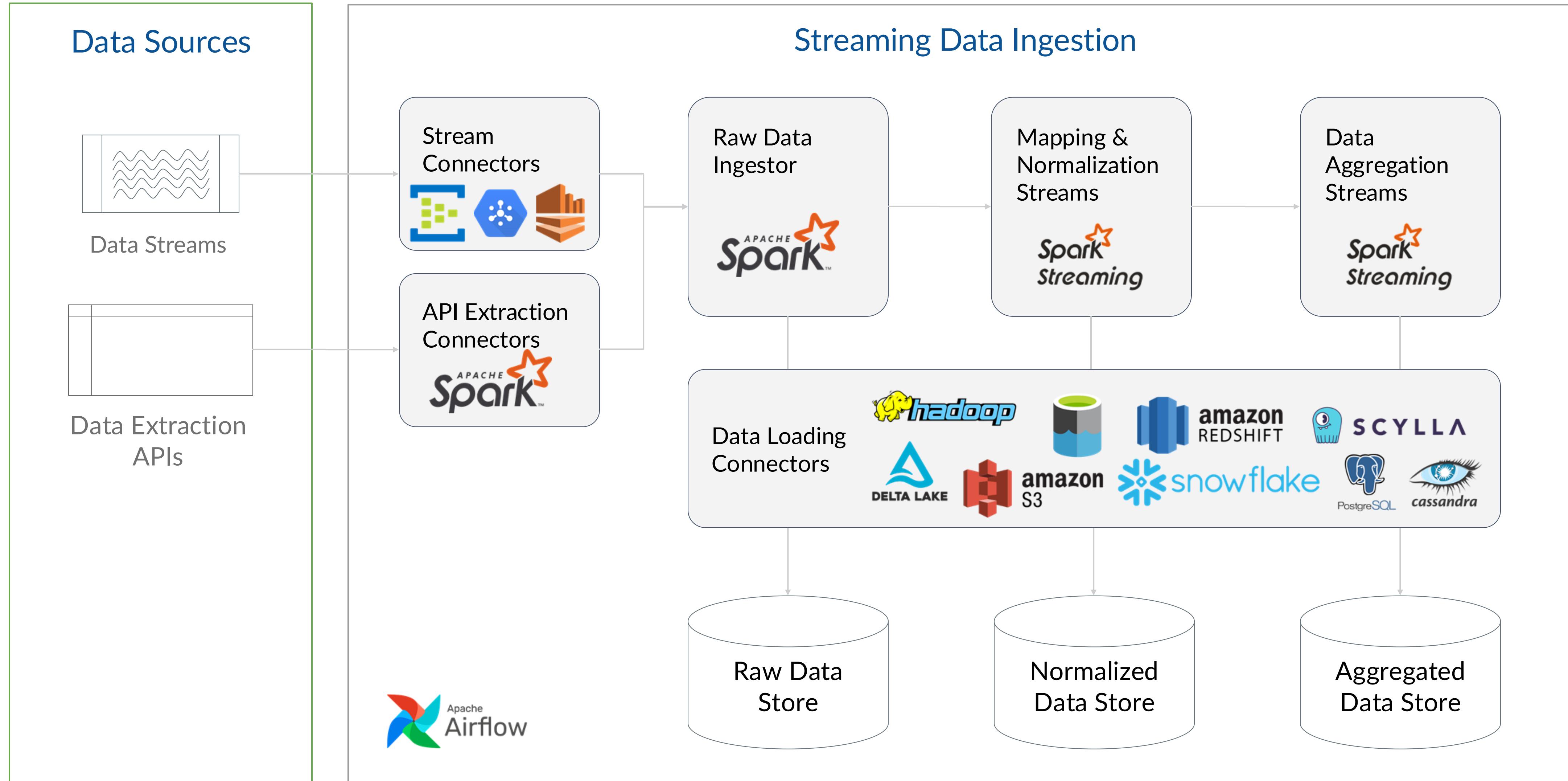


6. Electronic
approval



-  Digital Platform Development
-  Product Development
-  UI/UX Design
-  IoT Platform
-  Cloud-based SaaS Solution
-  Data Analytics
-  Machine Learning
-  Power BI Development
-  Azure Managed Services

Streaming Data Ingestion Data Flow



Case Study: Dynamic Pricing Analysis



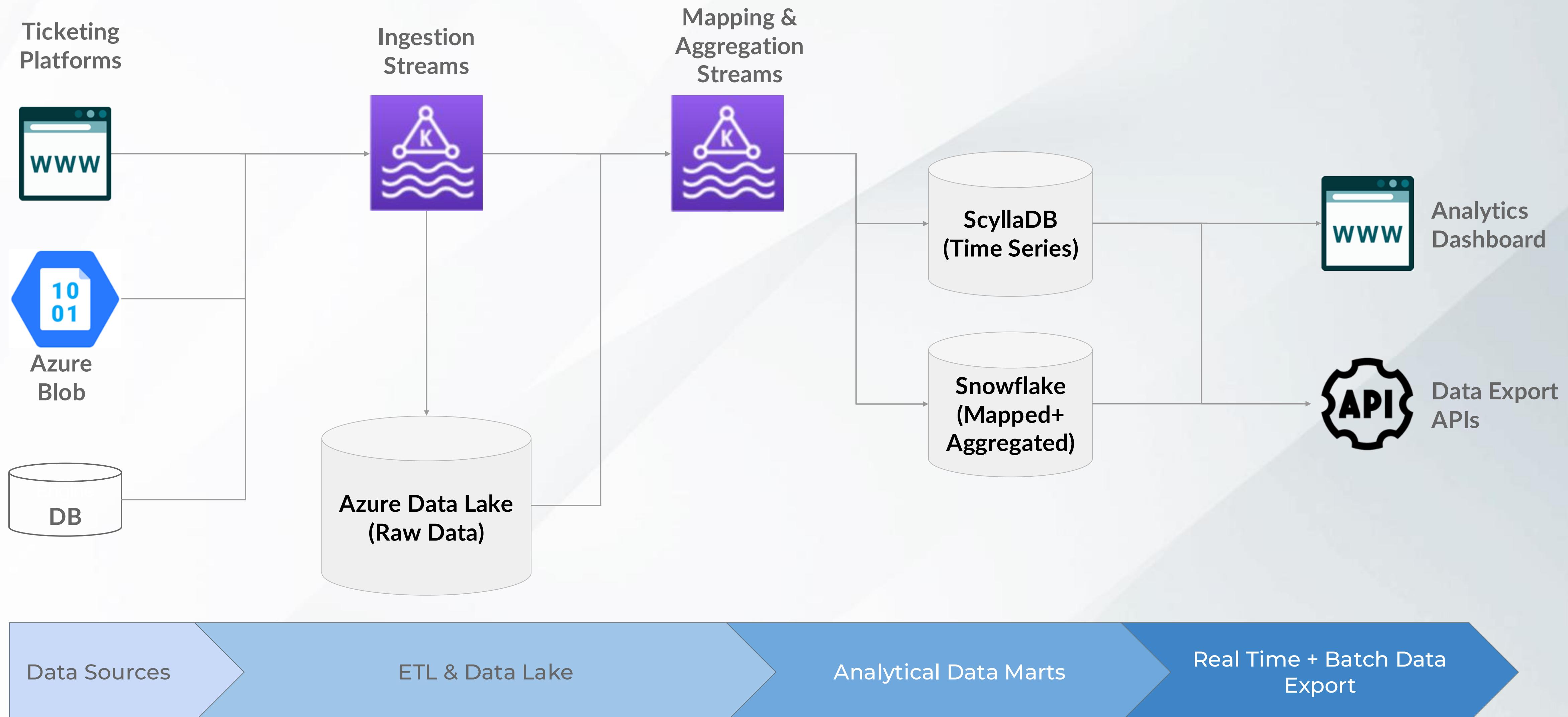
Requirements

- Platform ingests data from multiple ticketing engines to provide a dynamic pricing analysis to brokers
- Customers (brokers) require access to near real time streaming and batch analysis of pricing data
- Latency for pricing updates expected to be less than 1 minute
- Millions of data records per day to be ingested and processed, thousands of brokers accessing the platform using Analytics Dashboard as well as Export APIs

Solution

- Streaming ETL Engine using Spark on clusters with implicit data parallelism and fault tolerance
- Data Lake using Azure Data Lake (HDFS Parquet)
- Pricing Time Series Data in ScyllaDB and PostgreSQL containing mapped and aggregated data
- Orchestration using Airflow
- Customers consume near real-time data streams for various data sources per need
- “Pub/sub” model for data delivery, with available REST endpoint for data export

Data Flow



Case Study: AI based Audience Projections



Background

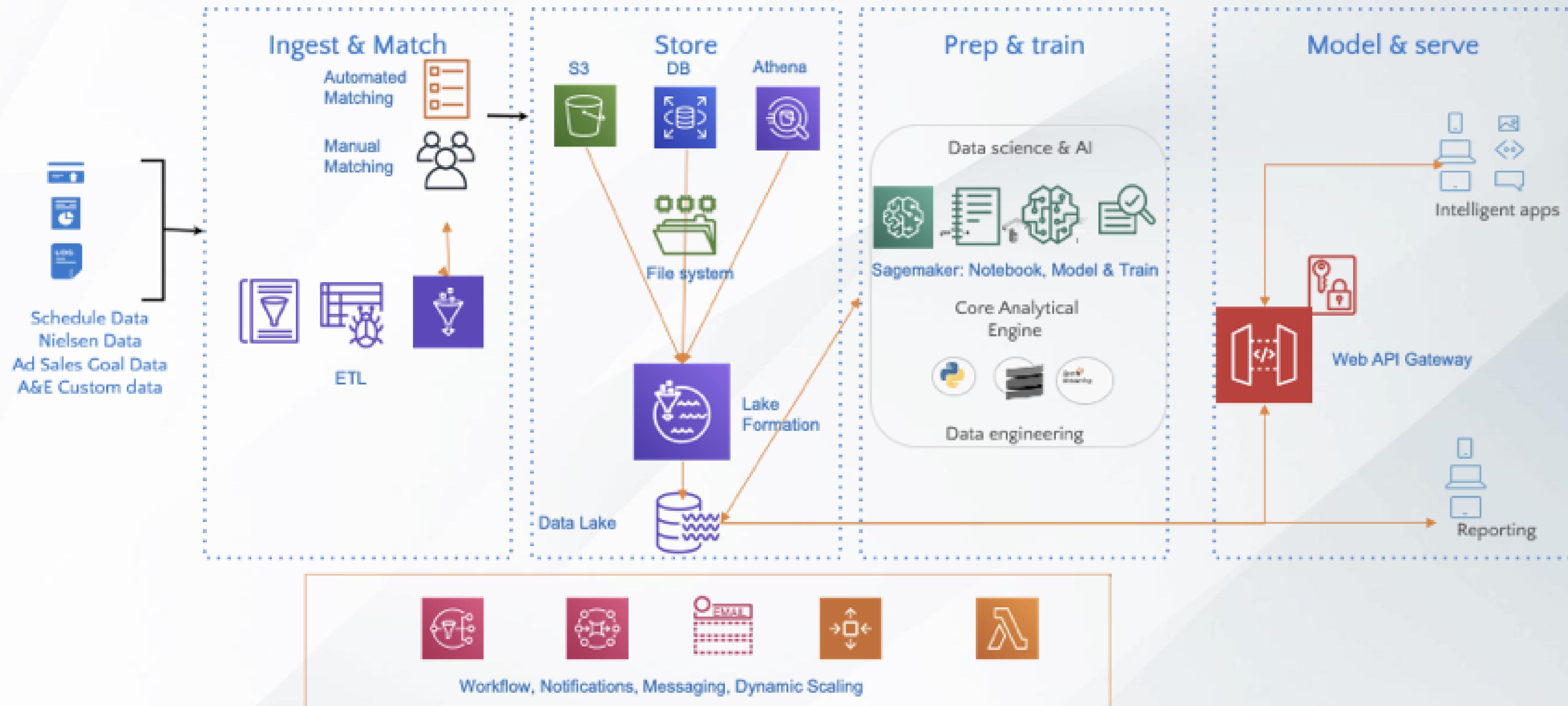
Client would like to attain the capability to perform audience forecasting in a consistent and reliable manner for various demographics for six networks so that they can:

- Reduce the manual efforts of maintaining their projected estimates/revenue
- Price their inventory better
- Incorporate updated estimates into ongoing program scheduling decisions
- Track Performance on a daily basis - planned vs actual for Audience delivery to further enable current and projected audience delivery vs. revenue goals.

Solution

- To perform data ingestion and matching of the diverse data sources to populate data lake based structures designed to support the AI/ML models built in the cloud
- The predictions of the models provide a point and a range estimate for audience projections which is then exposed through API
- The APIs serve a frontend application with a projections workflow that allows ad-sales analysts to consume, edit and finalize projections for their viewership and associated ad-revenue.

Architecture Diagram





Thank You

For more info please visit www.accionlabs.com