hoonartek

Delivering Databricks Driven-Modernization Programs

HoonartekRiverSide®

ISO 9001:2015 ISO 27001 ISO 22301 FSQS REGISTERED HELLIOS

Hoonartek - Empowering Enterprises

Liberating Data Democratizing AI

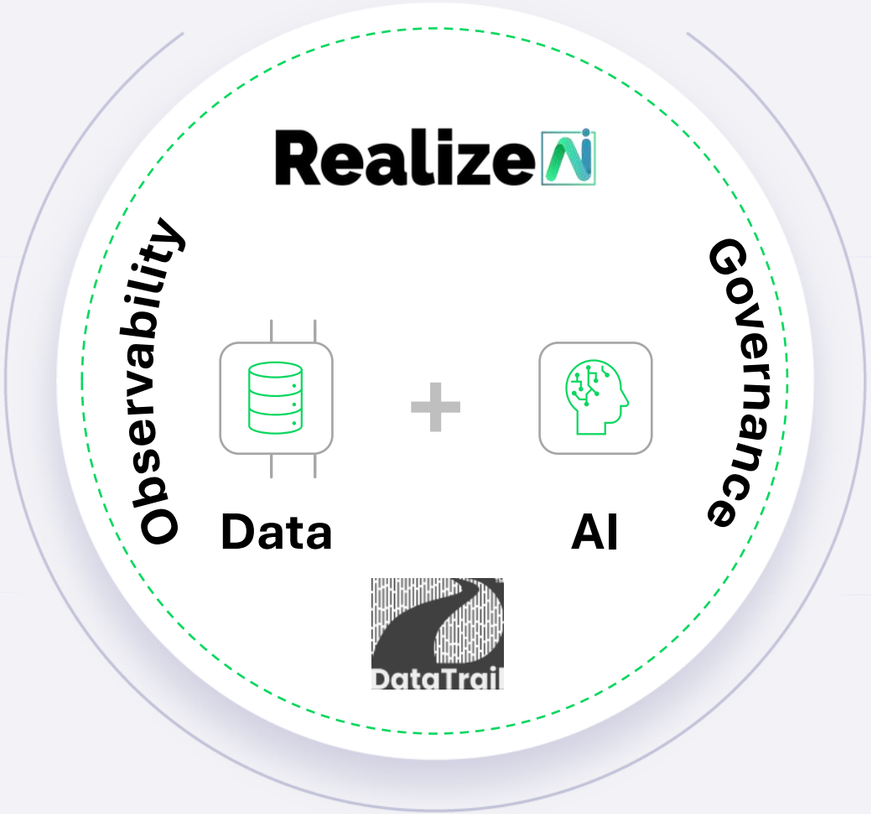
15+ years of Empowering users
Digital. Data. Analytics

Strategic Partnerships with technology leaders

Databricks, Snowflake, AWS, Azure,
GCP, Alation, Ab Initio, Experian

Unmatched AI & data engineering skills

~1000 expert consultants
driving transformative AI/ML solutions



Data & AI focused Global SI

USA . UK & Europe . Australia . ASEAN,
Backed by Certified Delivery Center in India

Solution Accelerators & Frameworks driven

Migration Factory for data migration
DataTrail for Governance and Lineage
Metadata driven CDC & ingestion toolkit

RealizeAI – Vertical AI & ML Solutions

BFSI, CMT, Healthcare and Manufacturing
AI Studio with 30+ pre built apps

Empowering Customers Globally

hoonartek

databricks

Migration Factory™

DataTrail®

RealizeAI

Data Engineering

- Streaming and batch ingestion framework
- Data quality & observability framework
- Fully automated deployments

Lakehouse

- Enterprise-wide subject area
- Logical Models for FSI and telecom
- Business Terms & definitions

Governance & Security

- Business glossary & catalog for FS, Telecom
- End to end lineage and governance setup
- PII and CDE discovery and tagging

Advanced Visualization

- Industry analytics pack for FS and Telecom
- KPI-driven insights for CFO Dashboard
- Semantic layer setup and metadata delivery

AI/ML & Generative AI

- 30+ readymade apps with Hoonartek AI studio
- RealizeAI framework for delivery of AI & ML use cases

40% faster and risk-free delivery with ready to deploy framework and prebuilt AI apps

RealizeAI

RealizeAI

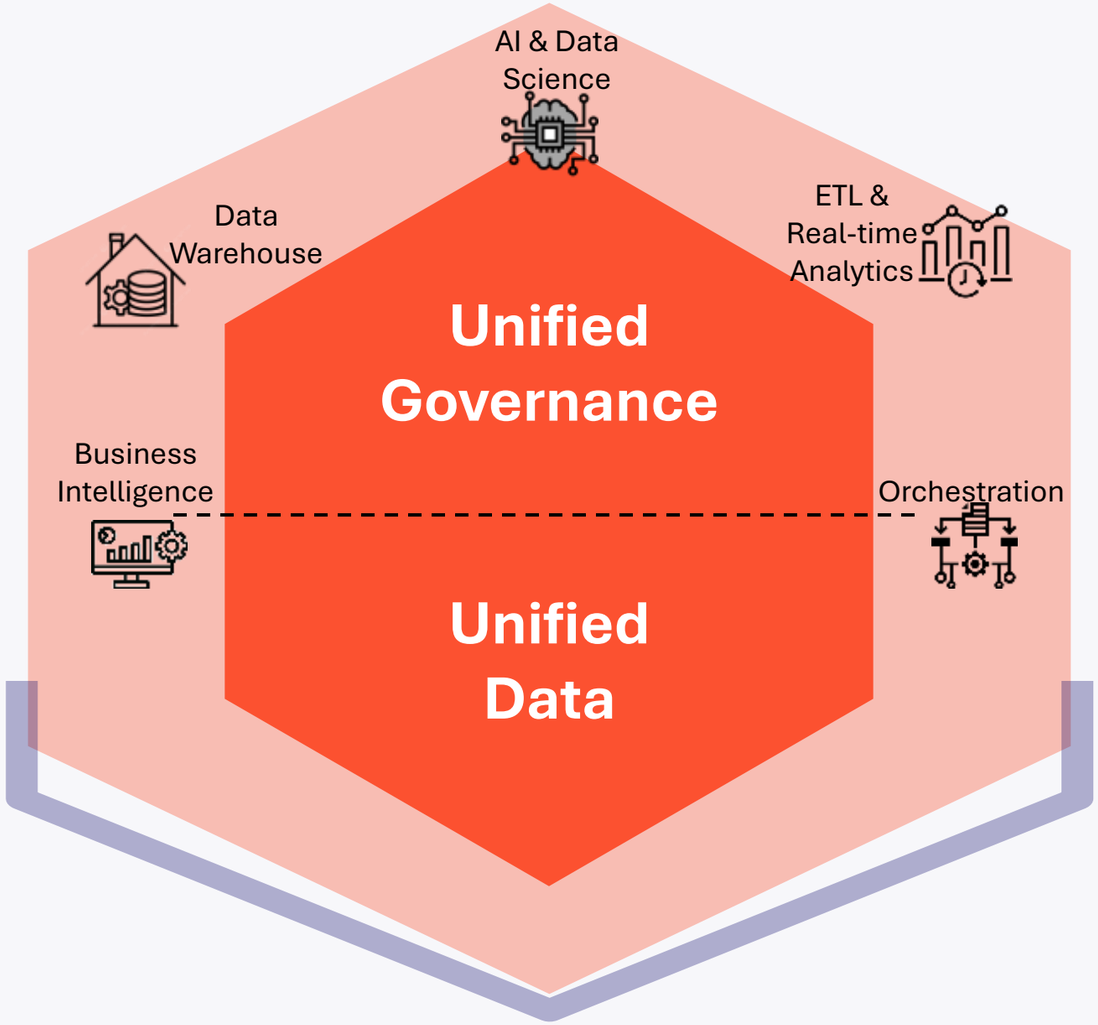
Data Intelligence Platform

Industries

BFSI, CMT, Healthcare, Manufacturing

Use Cases

Churn Prediction
Loan Approval Agent
Medical Image Analysis Agent
And more..



Business Value

Boosts customer service retention
Regulatory compliance

Powered by Databricks

Integration with Databricks for scalable ML workflows

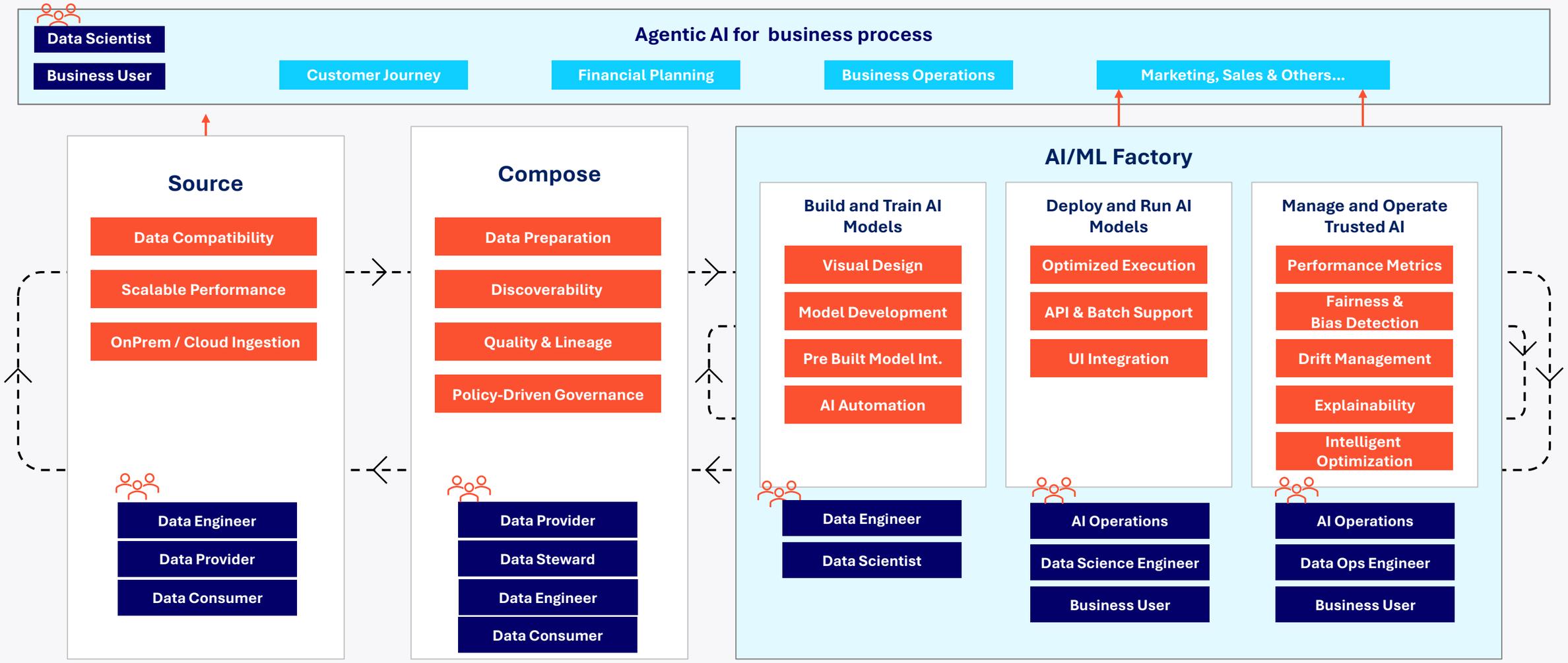
360 Capabilities

Vertical Agentic AI solutions
Predictive & generative intelligence
Domain-tuned model development
Rapid prototyping & deployment

Accelerated Delivery

Built by Hoonartek's AI Lab
Scalable and collaborative
Fast, efficient, end-to-end AI

Reduced time to market with AI ML Factory



RealizeAI Delivery Framework

AI Studio with 30+ Apps: Built by Hoonartek

Loyalty Management Platform
Auto-rewards customers in real time

Loan Delinquency Prediction
Identifies potential defaulters early

Medical Image Analysis
AI-assisted anomaly detection in scans

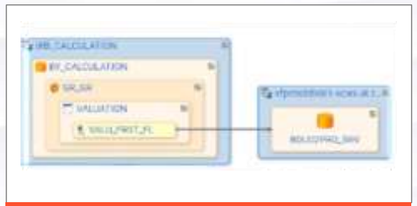
Fraud Detection
Flags suspicious credit card activity

Auto Claims Approval
Processes insurance claims autonomously

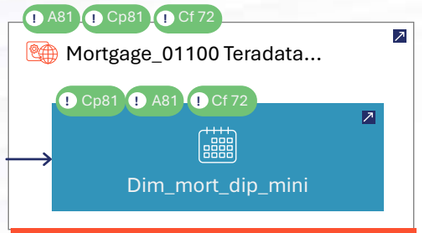
Churn Predictor & Retention
Detects churn risk and boosts retention

Data Governance

Enterprise Data Governance — Powered by Unity Catalog



Business Lineage



Data Quality over Lineage

Technical Lineage

Business Owner Olivia Bennett

Data Steward Thomas Hayes

SME Ankita Satpute

Accountable Stewards / Persons

40% faster compliance through automated classification & lineage (e.g., GDPR, DPDP)

25% risk reduction with clear ownership & stewardship

30% improved decision accuracy via trusted, high-quality data

Enhance **audit readiness by 45%** with automated metadata tracking & reporting

50% faster insights using integrated catalog & glossary

Up to 35% cost savings from centralized governance automation

Critical Data Element: Yes

Personally Identifiable Information: No

Field Value Derived: No

Sensitive: Yes

Classification

Customer MSISDN

Customer MSISDN _ Mobile Station International Subscriber Directory Number

Type: Business Term

Dictionary: Marketing Analytics

Business Glossary

General Information

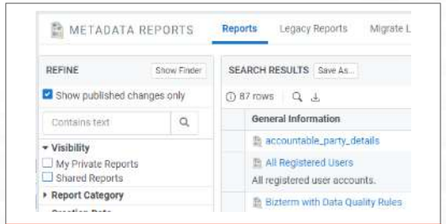
Table Column nk_policy_rules_passed_ind

Table dim_mort_dip_mini

Business Team Policy Rule Indicator

Indicator of policy rule passed (1=Yes, 0=No, 2= Don't know)

Linking Data Catalog and Business Glossary



Metadata Reports

Product Engineering: Data Products for Value Realization

Digital Engineering

- End-to-end product development and modernization
- Embedded agentic AI and ML Model for automation
- UX/UI with rapid prototyping
- Agile & DevOps delivery frameworks

Application Engineering

- Enterprise application modernization
- API-led integration & microservices architecture
- Custom software development
- Low-code / no-code platforms
- Continuous integration & deployment (CI/CD)

Cloud Engineering

- Cloud-native application development
- Cloud migration and modernization
- Infrastructure as Code (IaC)
- Kubernetes, containers & orchestration
- FinOps and cost optimization strategies

Data Engineering

- Scalable data pipelines (batch & real-time)
- Data lakehouse and warehouse architecture
- ETL/ELT modernization and orchestration
- DataOps and observability
- Data governance and quality engineering

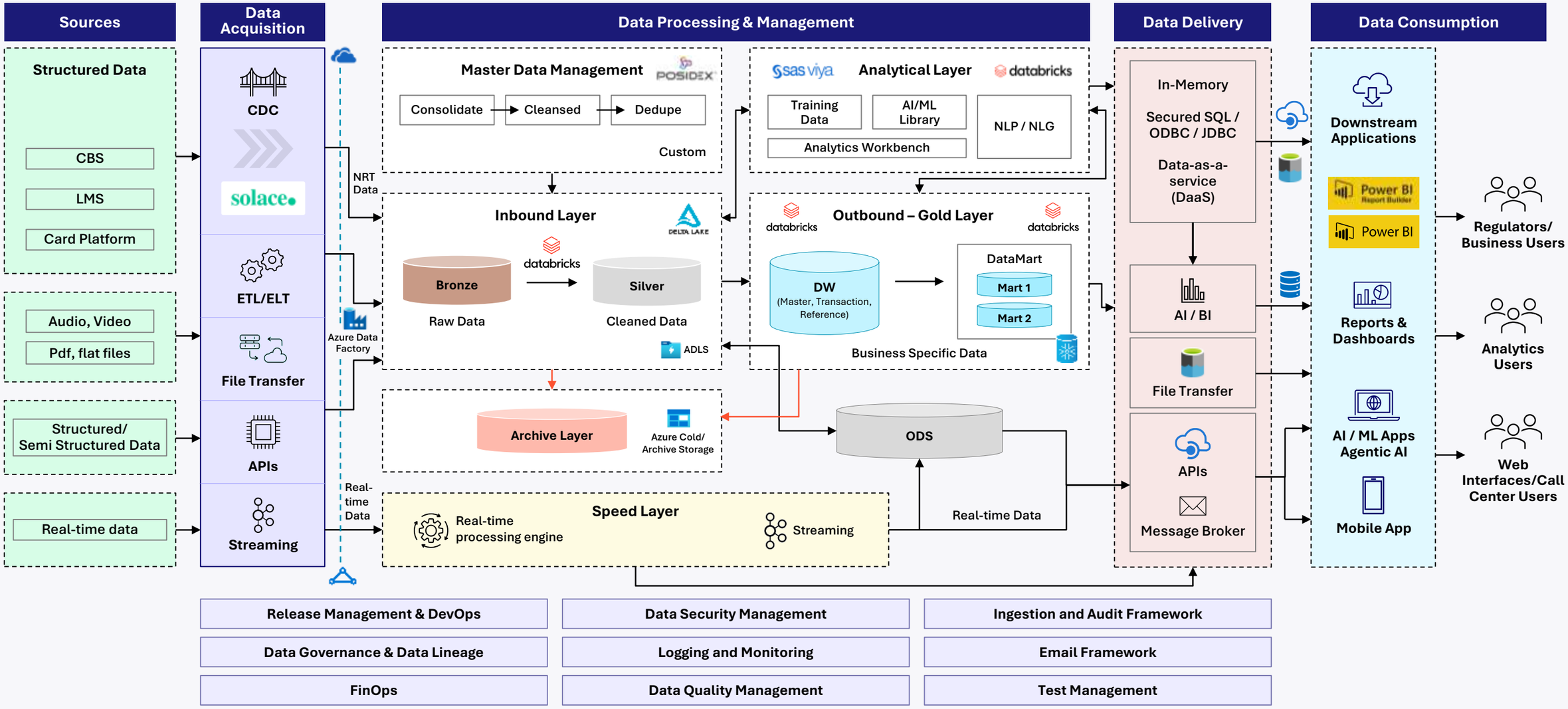
DevOps

Quality Assurance

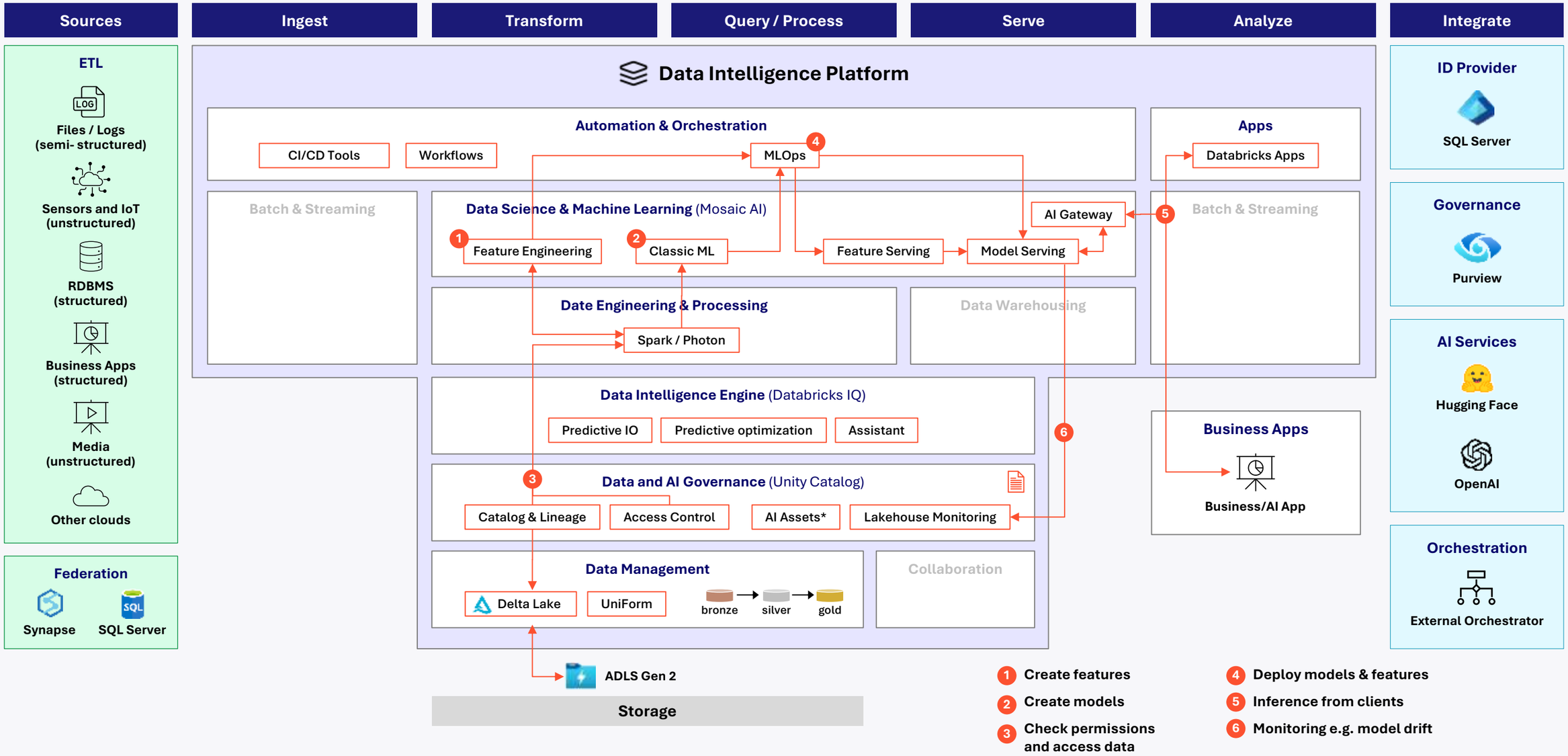
Technical Documentation

Managed Services

Blueprint: Financial Services Datalake

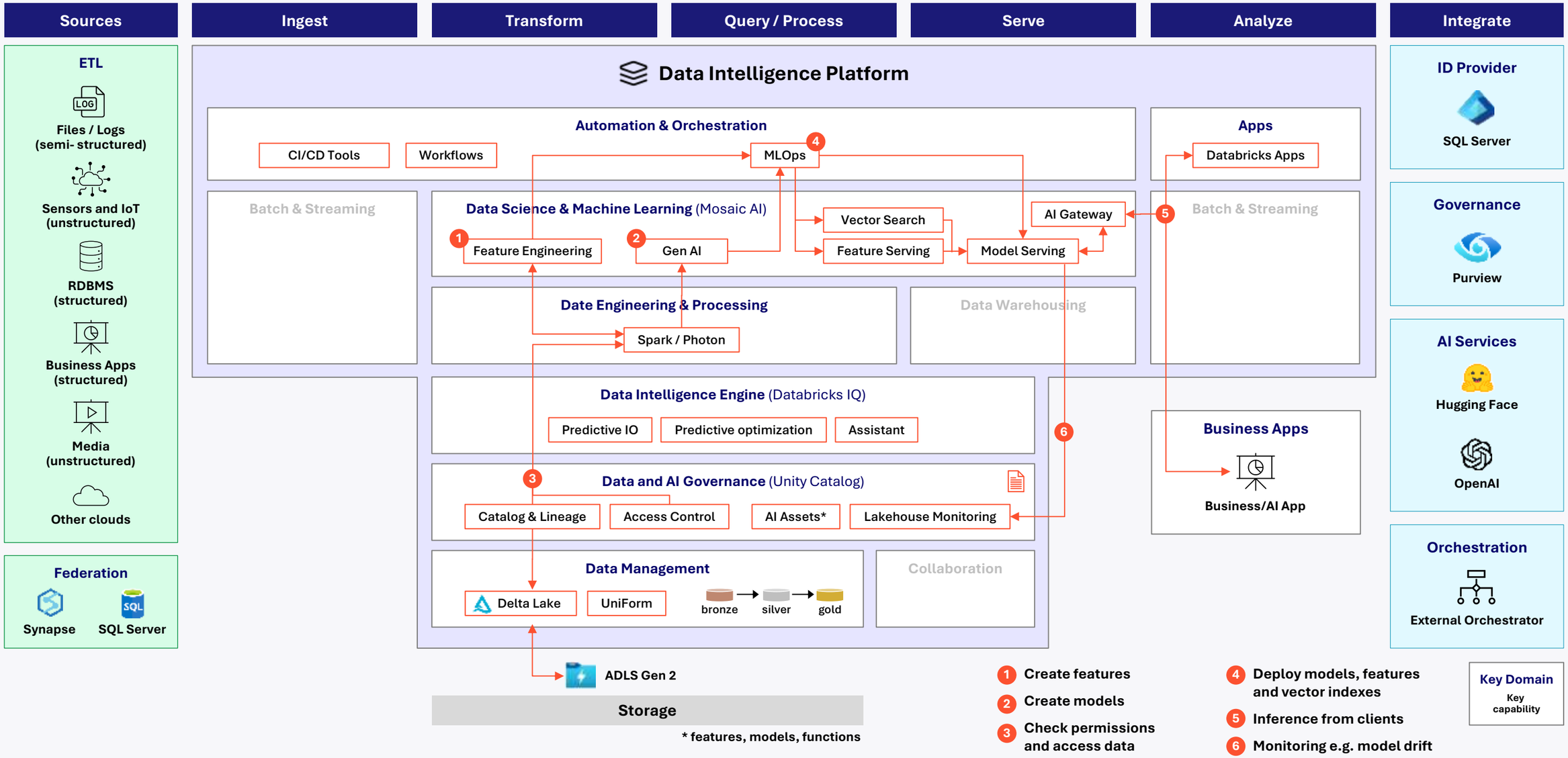


Blueprint: Financial Services AI / ML Apps

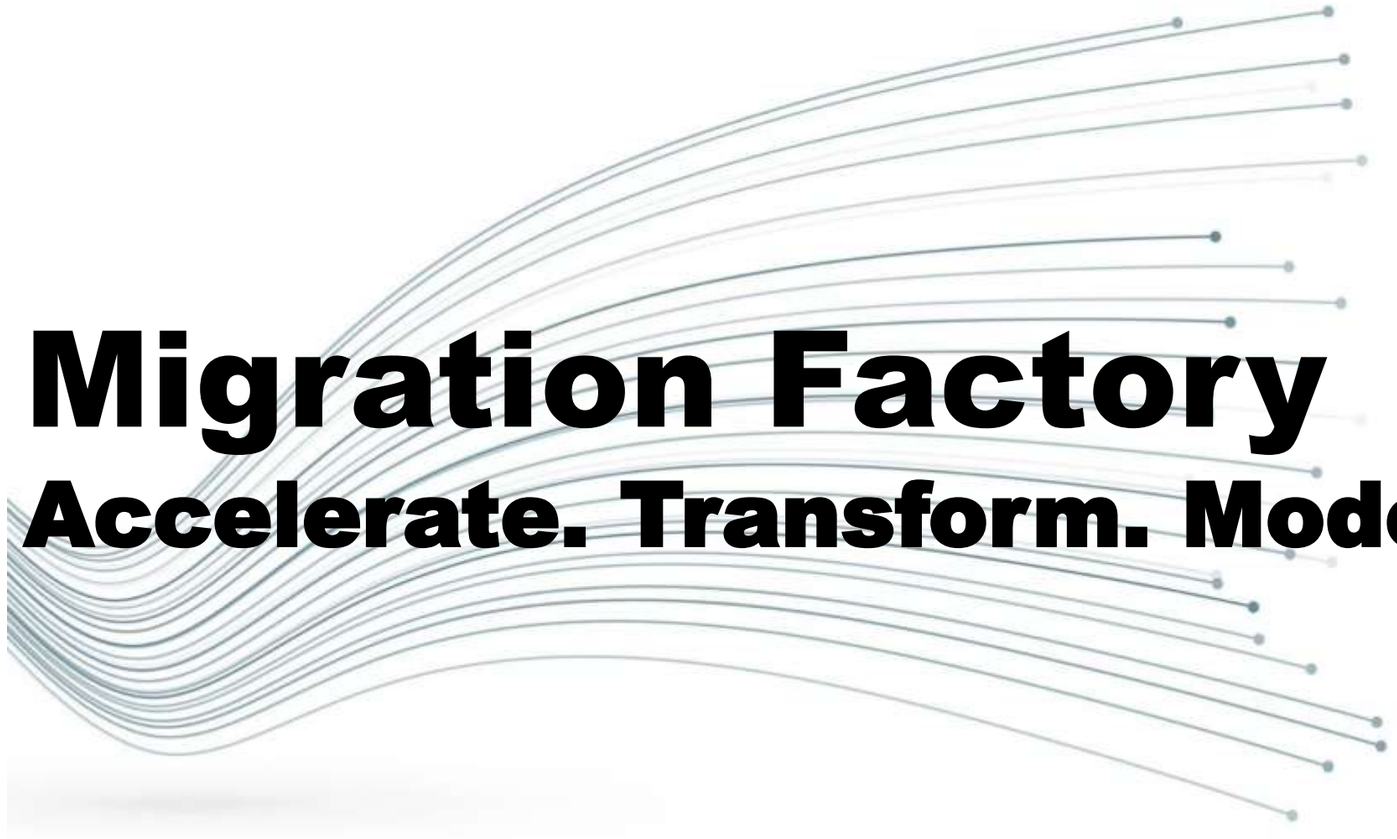


hoonartek databricks @ DAIS 2025

Blueprint: Financial Services GenAI Apps



hoonartek databricks @ DAIS 2025



Migration Factory

Accelerate. Transform. Modernize.

Accelerator Powered Agile Transformation

Scope & Requirement

- Plan & Estimator utility
- Current State Analyser

Data Migration

- Change Data Capture tool
- Split/Merge Utility
- Readymade Data Pipeline

Migration Factory

Build

- Iceberg toolkit
- AI Powered Code Converter
- SQL Validator

Data Quality

- Data Quality Framework
- DQ Reporting Dashboard

DataOps

- Data Catalog
- CostSentry
- Role Management Framework

Quality Assurance

- Automated Testing Framework
- Data Reconciler

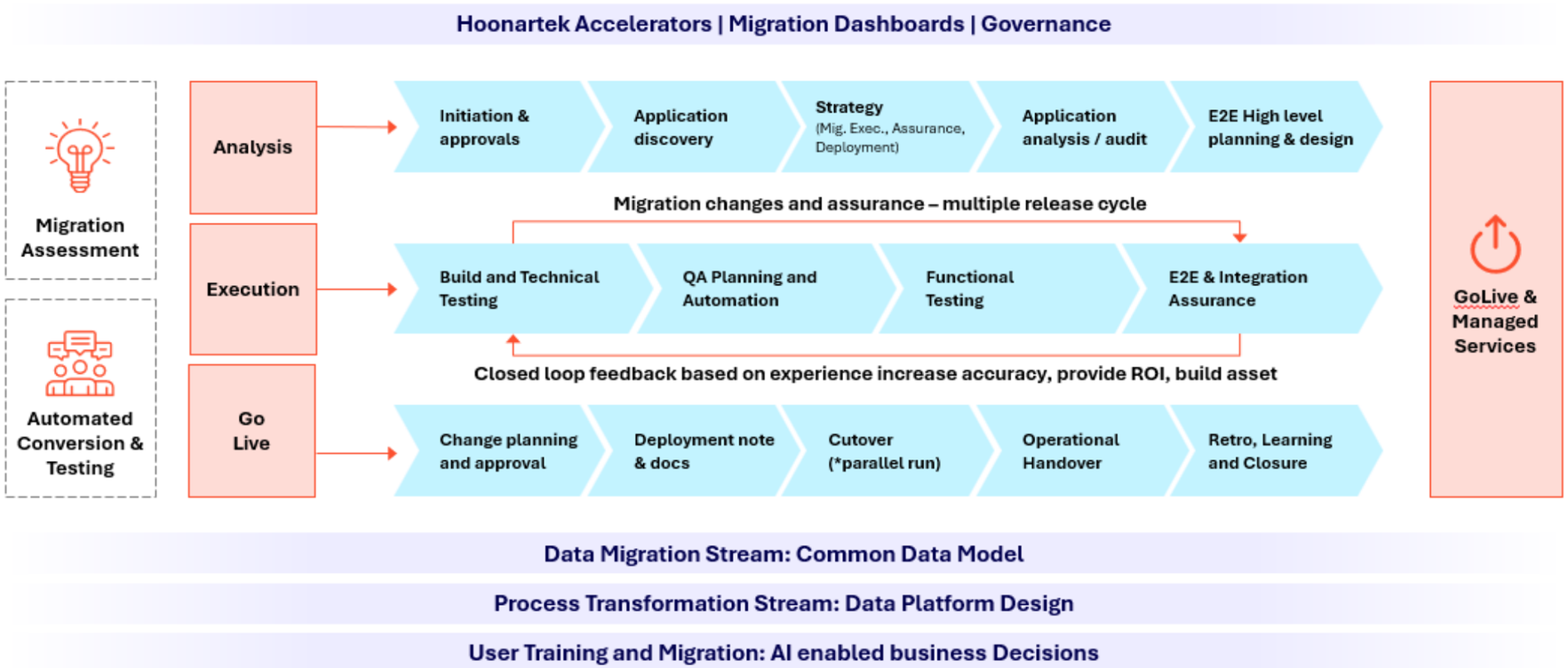
AI/ML & Data Sharing

- Peggy
- DataHaven
- Industry Model

Data Intelligence Platform

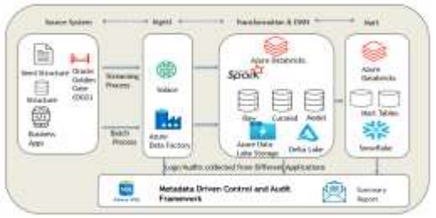
Migration Factory for risk-free transformation

IBM Datastage
mainframe
Greenplum
CLOUDERA
NETEZZA
ORACLE DATA INTEGRATOR
Spark SQL
Java
Synapse
Informatica
ORACLE
snowflake

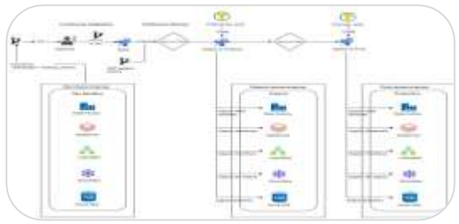


Data Intelligence Platform

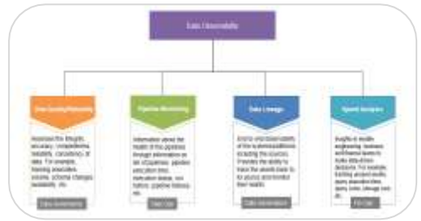
Reduced time to market with pre built accelerators



Ingestion and Audit Architecture Framework

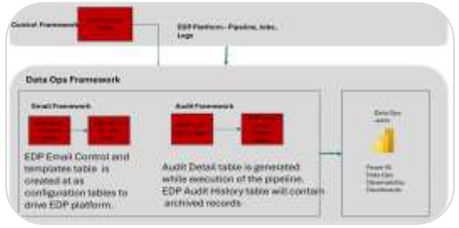


DevSecOps Framework

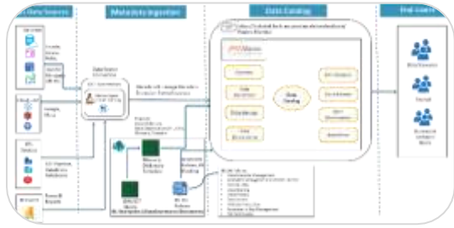


Data Observability Framework

Data Quality Framework

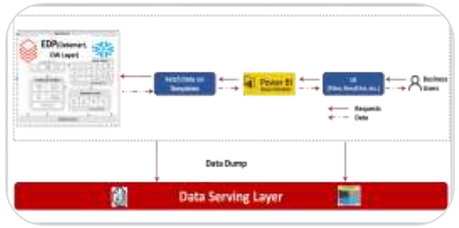


DataOps Framework



Data Governance Framework

FinOps Framework



Self Service BI Framework

Email Framework

hoonartek databricks @ DAIS 2025

DataTrail

Lineage . Compliance . Analysis

Data Privacy / GDPR Compliance Automation

Consent Management

Automated data lineage

PII Discovery

Policy Controls

PII Identification

Business Glossary

Masking Rules

PII Tagging

Data Catalog

DPDP/GDPR scope is complex, it cannot be done manually and kept in sync with source

Our DPDP/GDPR offering automates the collection of accurate evidence for regulatory submission

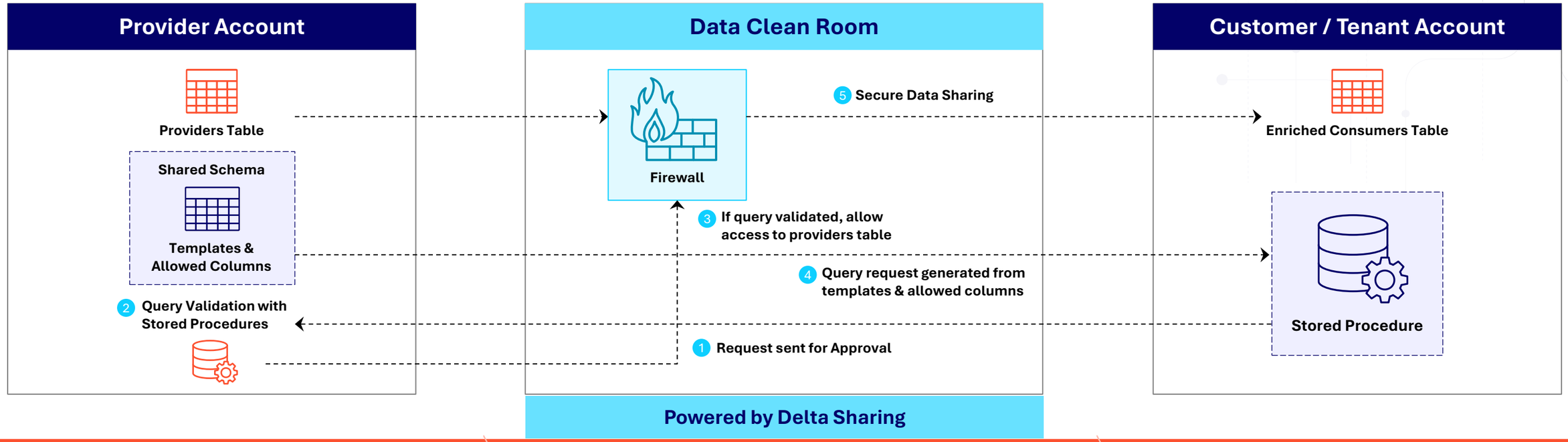
Our DPDP/GDPR offering ingests metadata “footprints in the sand” that were written each time data was processed by your IT systems

Predictable Databricks Smart GoLive Solutions

	MVP	Bootstrap	Startups	Hyper-Growth	A la Carte
Platform Fitment Evaluation	X	X	✓	✓	Requirement gathering
Architecture, Assessment, and Roadmap Definition	X	X	✓	✓	Scope refinement, finalization & approval
DDL Conversion	✓	✓	✓	✓	Assessment of current landscape
History Data migration & reconciliation	✓	✓	✓	✓	Solution development
Code conversion to Databricks SQL / PySpark	✓	✓	✓	✓	Project plan, estimation & timeline finalization
Unit testing	✓	✓	✓	✓	Solution Implementation
Zero to Databricks workshop	✓	✓	✓	✓	Consumption layer integration
Analytic, AI/ML, BI, Data sharing with Databricks	X	✓	✓	✓	Testing & business approval
Integration Testing and Data validation	X	X	X	✓	Parallel run & cutover
UAT, Parallel Run, and Cutover planning	X	X	X	✓	Managed Services
Databricks training	X	X	X	✓	Customized Pricing
Fixed Time | Fixed Outcome | Fixed Price **10x acceleration in customer go lives**		**Professional Services** • Solution Architecture • Consultation	• Data Migration • Advanced Training	• Data Governance • Data Quality	• Data Engineering • Managed Services
Typical Duration	2 weeks	4-10 weeks	4-10 weeks	10-16 weeks	Provided on Request

hcanartek databricks @ DAIS 2025

DataHaven: Data Exchange + Data Clean Room (DCR)



Secure & Scalable Data Sharing

Multi-Tenant Architecture
Scales businesses by adding group companies and advertisers as tenants

Intuitive Dashboards
Your data remains in your 'haven'

Intuitive Insights

Advanced Analytics
Automated reports made available at multiple levels

Itemized Bills
Monetize your data more granularly

Privacy

Access-Based Controls
Fine-tune access and usage in a privacy-preserving manner

Data-Driven FinOps: Insights, Optimization, and Efficiency

Analyze cost, cloud usage, and budgeting

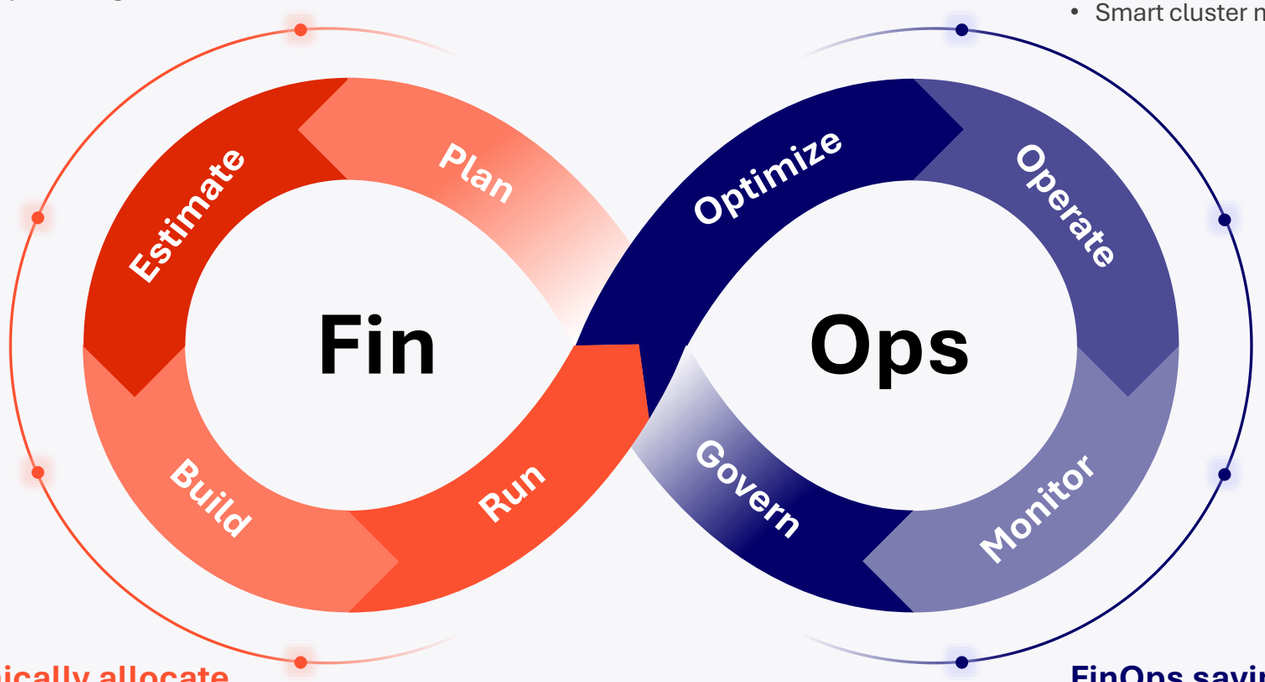
- Resource & Services Allocation
- Capacity Planning

Iterative optimize for cost/performance, Right sizing recommendations

- Autotuning
- Autoscaling
- Smart cluster management

Forecast and predict resources level usage

- Resource estimator
- Capacity Planning



Improve performance & efficiency

- Operational insights
- ML driven automation

Define SLA, job value and priority

- On-demand data profiling
- In-built data quality
- Schema drift detection & handling

Gain visibility into TCO

- Alerting
- Telemetry
- Elastic

Dynamically allocate resource to meet business goals, SLA

- Dynamic elasticity
- Cost Savings with GPUs
- Instance Selection

FinOps savings, reports, resource allocation, chargebacks, and enforcement

- Resource Dashboards
- Resource level Governance
- Open-metrics, Grafana

Proprietary & Confidential

Strategic alliances & partnerships

	Hoonartek and Microsoft deliver transformative solutions that drive innovation and operational excellence. Our partnership blends Microsoft's powerful platforms with Hoonartek's expertise in digital engineering, analytics, and cloud-native capabilities.
aws	Hoonartek and AWS enable high-performance engineering with cloud-native architectures, automation, and data-driven solutions. Our teams build resilient, scalable systems optimized for innovation and speed.
	Hoonartek and GCP help engineering teams build scalable, AI-driven cloud solutions that accelerate innovation and agility.
salesforce	Hoonartek and Salesforce empower businesses with tailored CRM solutions that drive efficiency and growth. Our experts ensure smooth implementation and long-term value.
	Hoonartek and ServiceNow enable the engineering teams to deliver intelligent digital experiences with automated workflows and advanced decisioning, driving agile, scalable, user-focused solutions.
	Hoonartek has a strong Snowflake Centre of Excellence NorthStar℠ with 50+ Snowflake consultants. With more than a decade of experience implementing turn-key solutions for some of the largest clients, We understands the value of data.
	Hoonartek and Databricks break data silos by bringing together data, analytics, and AI/ML seamlessly. Using open standard Databricks Platform, we bring agile data operations possible. Innovate faster with Hoonartek.
Ab Initio	Hoonartek is a Tier 1 full-service Ab Initio partner, known for delivering complex, high-impact projects globally with a 100% success rate. Our dedicated innovation lab—first of its kind—keeps us at the forefront of Ab Initio excellence and best practices.

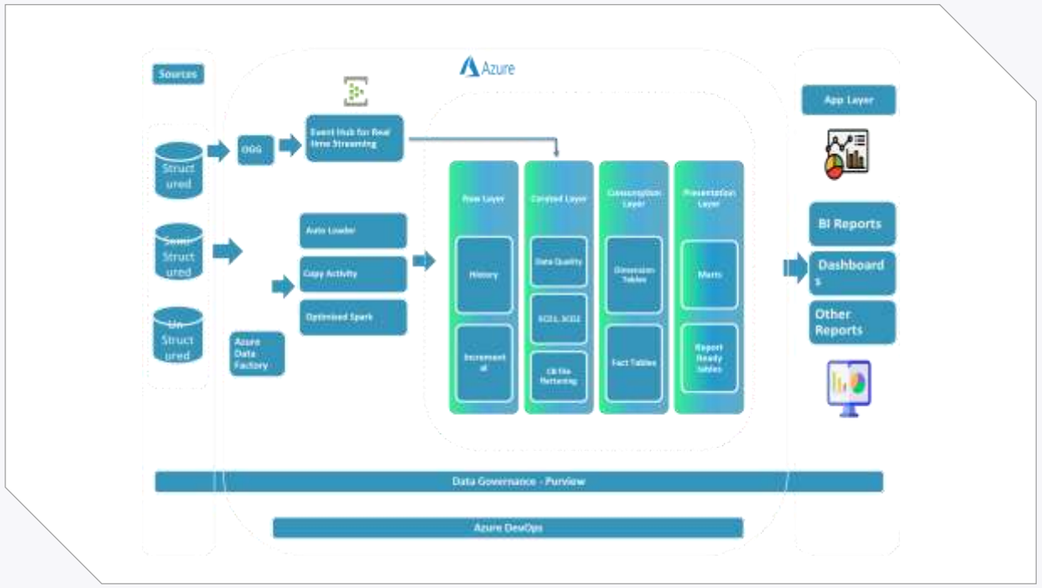
Data Platform Modernization for a Leading Bank

Business Problem

- Providing real-time disbursement and collections data for the 16 million customer from 157000 locations
- High concurrent volumes of logged in users accessing the real-time operational insights and batch solutions
- Data from POS / mobile disbursement and collections
- Legacy data platform on premise, with scalability and data quality issues and high TAT
- AI / ML capability non-existent and reliance on traditional Excel / BI report driven reporting / KPI
- Evolved architecture and organic data model which is not architected

Solution

- **Modern data platform** implemented on databricks platform with ingestion, data governance, fin ops and certified data quality
- Data catalog and lineage accelerated data products rollout and consumption
- AI/ML insights powered segmentation, nudges, risk, and lifecycle management
- Real-time data APIs provided on-demand data delivery and an insights layer
- Regulatory compliance ensured seamless RegTech extracts for bureau and central bank submissions/audits



Value Delivered

Data Product Launches

- 3x improvement in number of products

Faster time to Market

- 35% reduction in average data product launch

User Concurrency

- 2000 power user
- 150,000 static published dashboards

Governance and Compliance

- 100% Data Catalog and Lineage with certified data

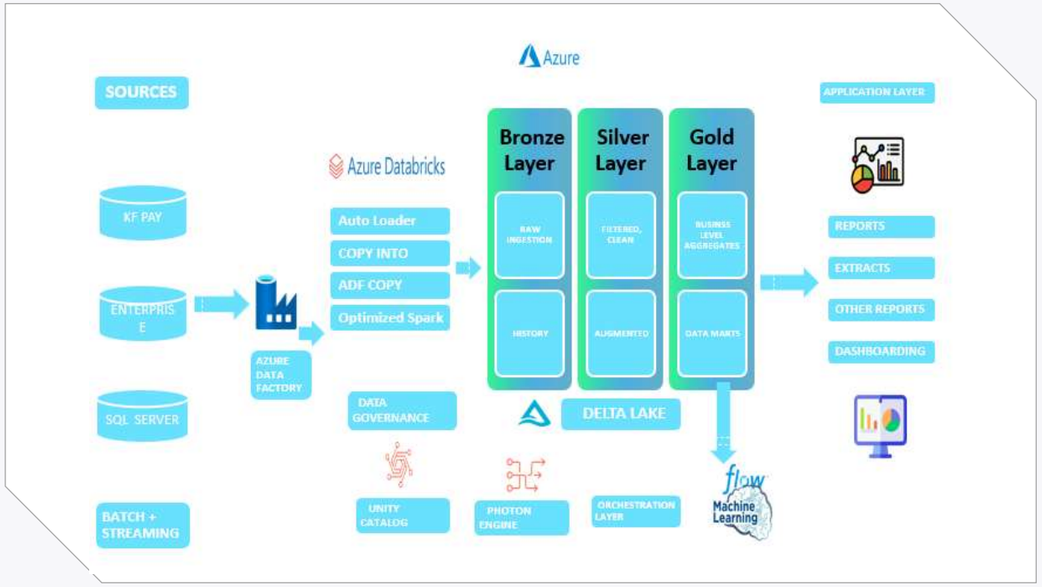
Data Products Delivery for HR Consulting major

Business Problem

- A global HR consulting and data products company wanted to transform their customer experience
- Legacy product stack was built on the traditional ETL and data warehouse
- The business journeys did not have embedded insights in the colleague and customer facing apps
- The data discovery governance for understanding lineage and data accuracy
- AI / ML capability non-existent and reliance on traditional Excel / BI report driven reporting / KPI
- Redundant data products with inconsistent information and manual processes

Solution

- **Revamped** architecture with replacing the silo'ed data islands into a medallion architecture with rationalized data products
- Data catalog driven data engineering framework covering CDC, ingestion and data platform using delta tables
- AI / ML platform to cover the entire MLOps lifeycle from model building, serving, inferencing for embedding it into business functions
- Integration of the new platform with automation and finops for reduced TCO
- Data literacy program roll out for business for adoption of the new platform



Value Delivered

Automation for reduced TCO

- 35% reduction in the support budget with automation and self healing

Data culture adoption

- Increase of the userbase of data products by 5x

AI ML Embedded insights

- Transformation of 12 key business processes with embedded AI
- Provisioning of ~20+ data APIs for the faster time to market

Data Trust

- Consolidation of 8 data silos into one holistic data platform
- 10x reduction in the number of user queries for manual data pulls

Loyalty Management platform with AI Agents

Business Problem

Interlocking business challenges that surface when a manual-receipt process is used to issue loyalty cards :

i. High Error Rates

ii. Operational Overhead

iii. Poor Customer Experience

iv. Missed Analytics & Personalization

Value Delivered

Data Driven Insights

30% higher signup rate

Faster Enrollment

80% faster Enrolment

Cost savings

50 % reduction in labour costs

Enhanced CX

25% lift in repeat visits

Improved Accuracy

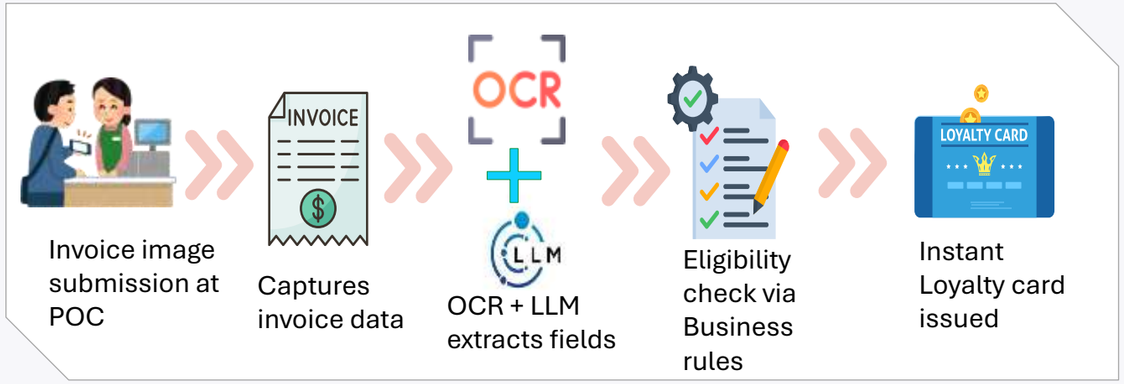
95% field-extraction accuracy

Solution

Agentic workflow driven solution

1. **Customer submits** invoice at POS or uploads image in the mobile app
2. **Loyalty Agent c**aptures invoice image and forwards to its AI service LLM Invoice Parser – Uses OCR + LLM to extract fields (invoice No, date, line-items, totals)
3. This data is sent to the **Loyalty Application** that Consumes the data, applies business rules (e.g., spend thresholds) to evaluate eligibility
4. If eligible, generates digital or physical **loyalty card instantly**, managing personalisation and wallet provisioning
5. API to CRM / Marketing Systems captures enrolment events for campaigns, analytics, and follow-up offers

Process Diagram



Churn Predictor & Retention AI Agent

Business Problem

Unexpected customer attrition with no proactive mechanism to detect or prevent churn.

I. No early warning system

II. Revenue leakage from preventable churn

III. Inefficient retention campaigns

IV. Lack of insight into churn drivers

Value Delivered

Proactive Retention

10–20% reduced churn

Operational Efficiency Gains

40-60% reduced workload

Increased Lifetime Value (LTV)

5–15% increase in LTV

Data-Driven Strategic Insights

20–35% reduction in churn incidents

Higher Marketing ROI

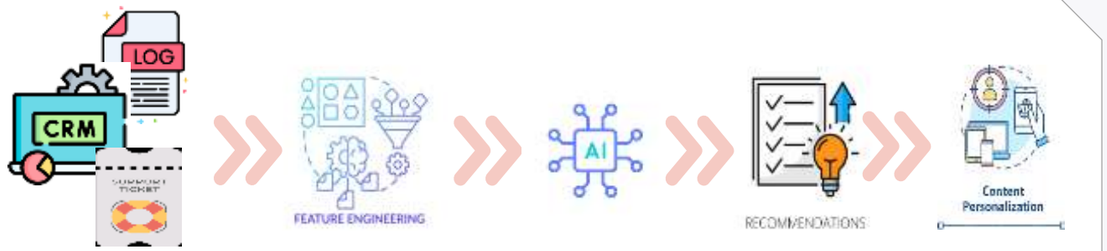
~30% better ROI

Solution

Agentic workflow:

1. **Data Ingestion**: Aggregates data from CRM, usage logs, and support tickets
2. **Feature Engineering:** Derives predictive signals (e.g., usage trends, support flags)
3. **AI Model:** Uses Gradient-Boosted Trees to predict churn risk
4. **API Output:** Returns churn score, top risk factors, and recommended actions
5. **CRM Integration:** Triggers personalized retention offers and logs

Process Diagram



Data ingestion from multiple sources | Derive trends/ usage | AI model predicts churn | Recommended actions | Personalized offers

Loan Delinquency Predicting Agent

Business Problem

Traditional methods of loan health prediction are found wanting because

1. Rising Loan Defaults
2. Inadequate Risk Models
3. Manual Underwriting Bottlenecks
4. Late Risk Detection
5. Regulatory Compliance Pressure
6. Poor Customer Retention

Loan health Monitoring is of utmost importance for the banking system to avoid Sub-Prime Mortgage Crisis like situation

Solution

- ML-based risk scoring engine to predict likelihood of delinquency at borrower level.
- Leverages transactional behavior, repayment history, credit utilization, income patterns, and demographics.
- Uses supervised learning models like Logistic Regression, XGBoost, and LSTM for time-based trends.
- Real-time scoring pipeline built on Databricks + MLflow, integrated with loan origination and servicing systems.
- Risk scores used to trigger proactive outreach, restructuring offers, or credit limit tuning.
- Feedback loop with collections data for continuous model refinement.

Value Delivered

Risk Mitigation

- 30–50% reduction in delinquent loan exposure through early interventions

Operational Efficiency

- 40% reduction in manual risk assessments and prioritization effort

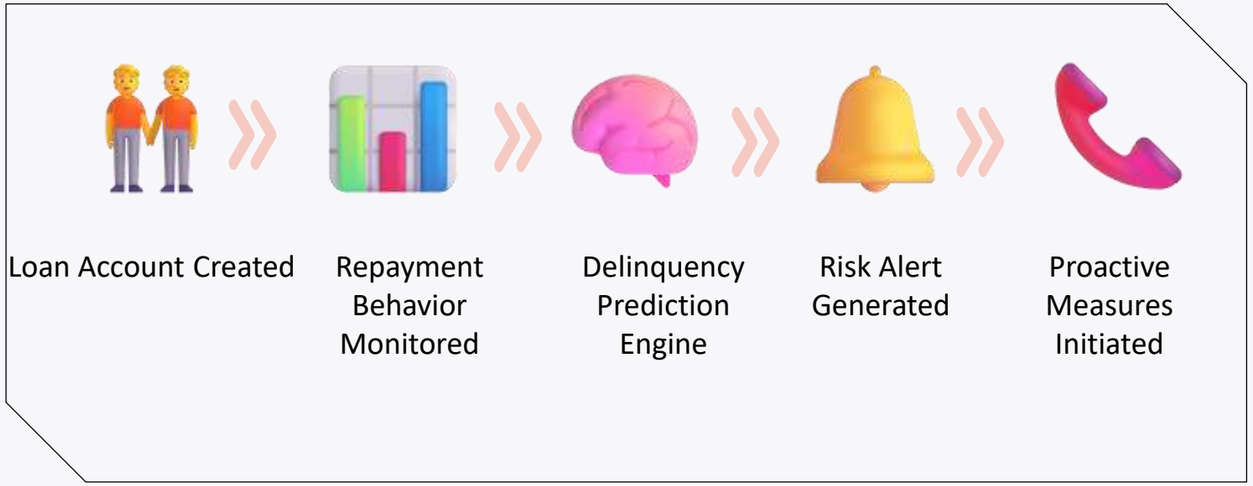
Customer Retention

- Improved borrower experience through pre-emptive engagement

Profitability Boost

- Improved NPA forecasting = optimized provisions and higher margins

Process Diagram



Medical Image Analysis Agent

Business Problem

Manual interpretation of medical images (e.g., X-rays, MRIs, CT scans) is:

I. Time-intensive
II. Prone to human error
III. Limited by specialist availability
IV. Causes diagnostic delays and variability

Solution

1. **Image Input** – Clinician uploads image file.
2. **AI Engine** – Uses computer vision + LLM to detect patterns, classify conditions, and highlight anomalies (e.g., tumors, fractures, lesions).
3. **Result Interpretation** – Generates structured report with annotated images and confidence scores
4. **Clinical Review** – Radiologist reviews AI findings, confirms diagnosis, and finalizes report.
5. **Integration** – Outputs fed into relevant systems, alerting clinicians and supporting treatment planning

Value Delivered

Faster Diagnosis

60% reduction in image review time

Operational Efficiency

50–70% reduction in interpretation time

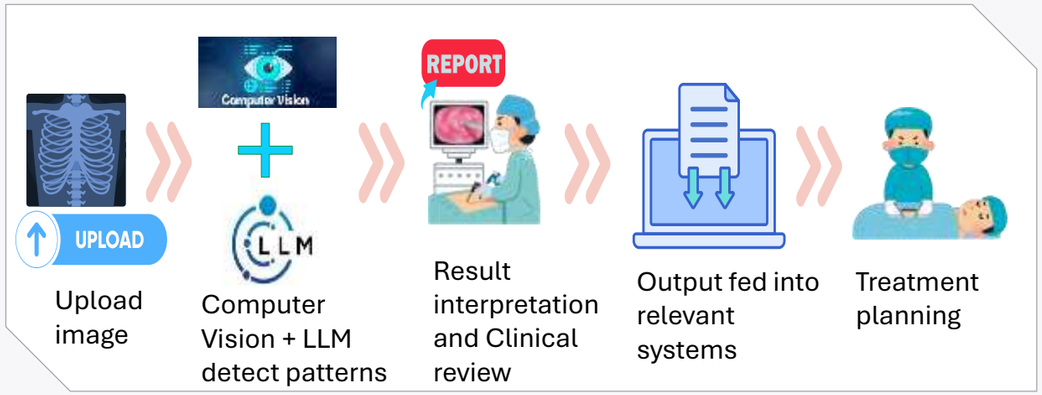
Time and Cost Savings

~80% reduction in manual efforts

Improved Accuracy

>90% diagnostic precision

Process Diagram



hoonartek databricks @ DAIS 2025

Autonomous Loan Approval Agent

Business Problem

Traditional methods of loan approval are found wanting because of

1. Lengthy Approval Cycles
2. Inconsistent Underwriting Decisions
3. Low Approval Rates for New-to-Credit Applicants
4. Poor Scalability
5. High Operational Costs
6. Compliance & Audit Gaps

Solution

- Deploy an Agentic AI co-pilot to automate loan decisioning with human-in-the-loop for edge cases.
- Ingests and analyzes application data, credit reports, income proof, bank statements, and alternative data.
- Uses ensemble ML models for credit scoring and risk categorization, backed by explainable AI (XAI).
- Performs document verification, fraud checks, eligibility evaluation, and decision recommendation in real time.
- Built on Databricks workflows, integrated with CRM, core banking, and API gateways.

Value Delivered

Faster Loan Processing

- 70–80% reduction in approval turnaround time (from days to minutes)

Higher Approval Rates

- 25–30% increase, especially for new-to-credit and thin-file customers

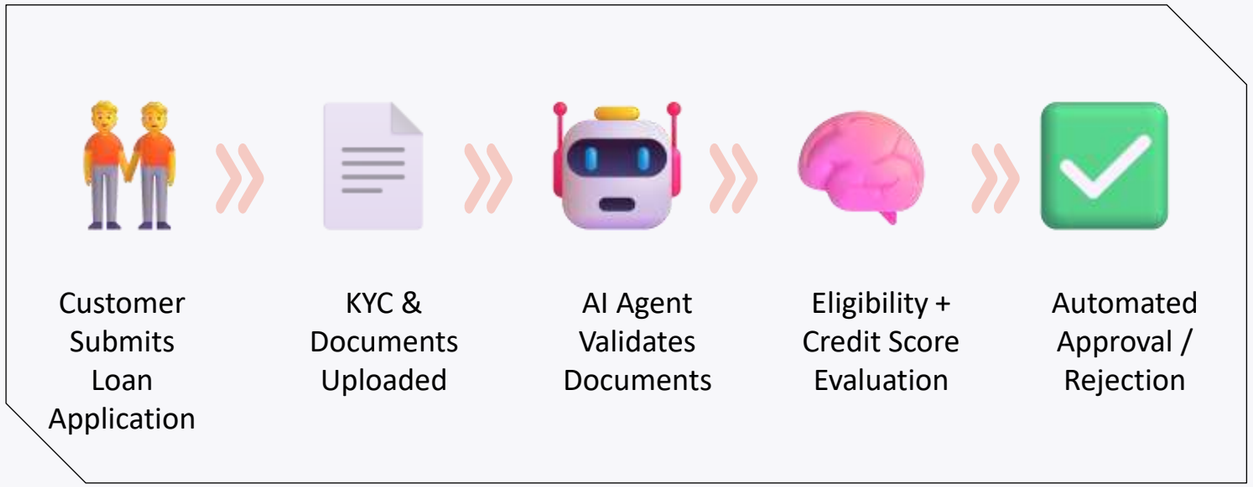
Risk Accuracy

- Improved loan portfolio quality with consistent, data-driven decisions

Scalability & Resilience

- AI agents can process 10x more applications during high-demand periods

Process Diagram



Proprietary & Confidential

Agentic AI-Automatic Claims Approval Agent

Business Problem

Why an automatic claims approval agent is the need of the hour?

1. Delays in Claims Processing
2. High Operational Costs
3. Inconsistent Claim Decisions
4. Rising Fraudulent Claims
5. Low Straight-Through Processing (STP)
6. Regulatory Scrutiny

Solution

•Agentic AI system automates end-to-end claims triage and decisioning.

•Ingests claim forms, images, supporting documents, and policy data via APIs and OCR/NLP.

•ML models assess claim validity, policy coverage, fraud risk, and auto-approve low-risk cases.

•Integrates with core insurance platforms, document management systems, and audit trails.

•Uses Explainable AI (XAI) to provide reasoning for approvals/denials.

•Human agents are only alerted for exceptions, borderline, or high-risk cases.

Value Delivered

Claims Turnaround Time

- 70–90% reduction — minutes instead of days/weeks

Cost Efficiency

- 50–60% lower processing costs by reducing manual workload

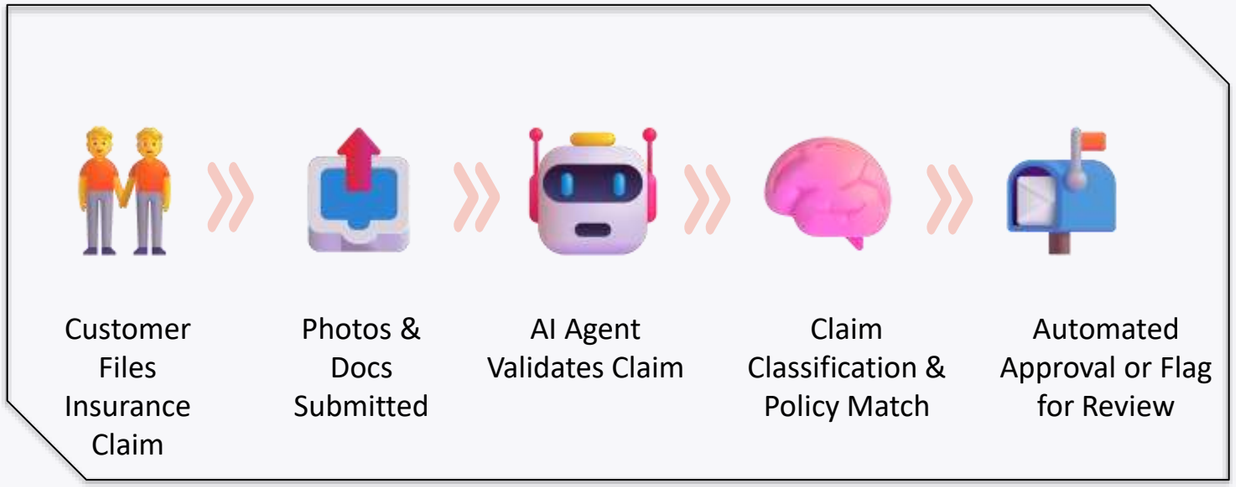
Compliance & Auditability

- Transparent decision rationale with explainable AI modules

Scalability

- 10x claims handling capacity during peak demand (e.g., natural disasters)

Process Diagram



Credit Card Fraud Detection

Business Problem

Why current methods of fraud detection are found wanting?

1. Evolving Fraud Tactics
2. High False Positives
3. Slow Response Times
4. Manual Investigation Overload
5. Compliance Pressure
6. Customer Trust at Risk

In 2023 USD 33.83 Billion were lost in Credit Card Fraud across the world.

Solution

- ML-powered real-time fraud detection engine integrated with core banking systems.
- Historical and real-time transaction data used to train behavioral and anomaly detection models.
- Algorithms: XGBoost, Isolation Forest, LSTM for temporal patterns.
- Real-time API deployment with scoring pipeline using Databricks + MLflow.
- Feedback loop for continuous model retraining and adaptive learning.

Value Delivered

Reduced Cost of Operations

- ↓ 30 % on fraud operations

Fraud Detection Rate

- ↑ 92–97% accuracy

Response Time

- Real Time

False Positives

- ↓ 40–60% reduction

Process Diagram

Customer Initiates Transaction → Transaction Data Captured → Feature Extraction by AI Engine → Fraud Detection Classification Engine → Alert Raised / Transaction Blocked