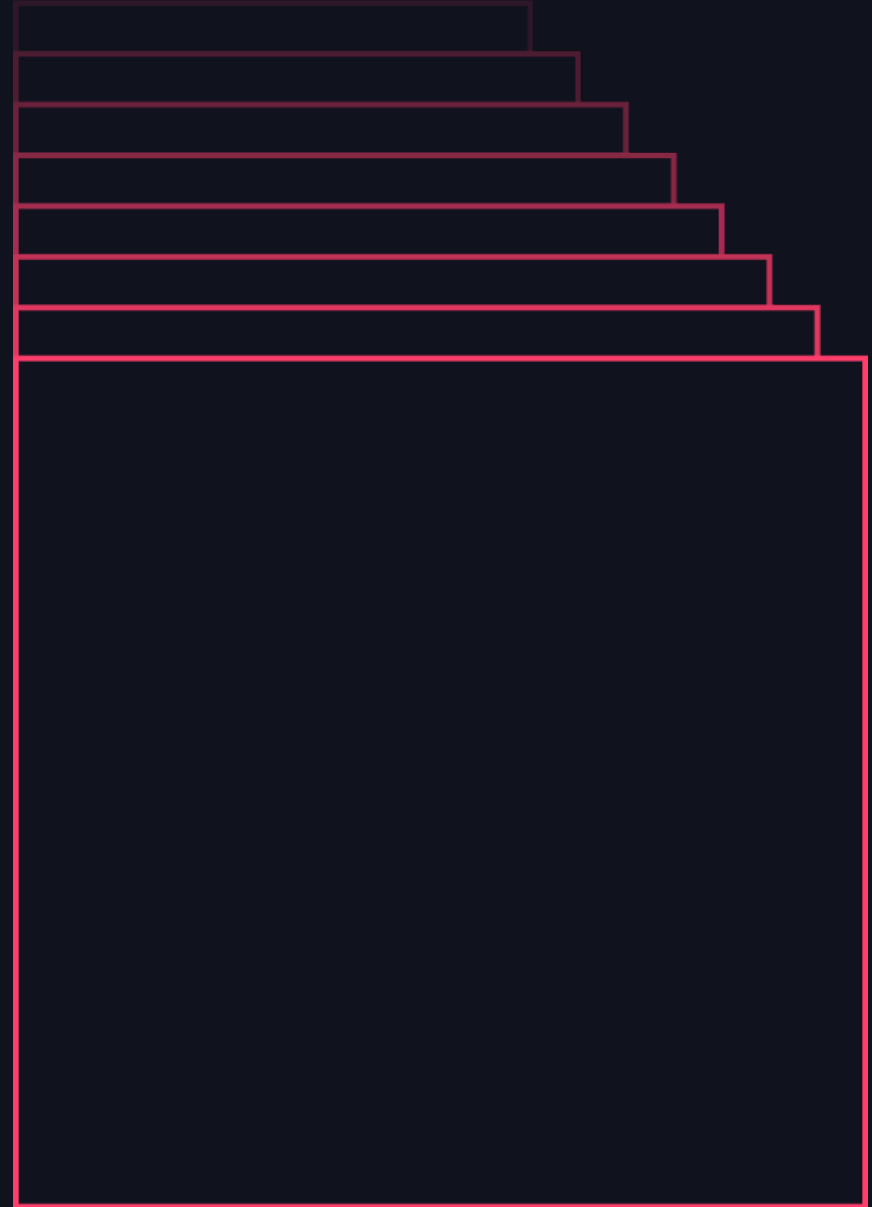# DATA+AI SUMMIT
BY databricks

Informatica|Enhance Enterprise Trusted Data Governance Integrated w/Databricks SQL UC

Pradeep Anandapu & Ajay Gollapalli
Date 11th June 2024

Today's Presenters

Pradeep Anandapu

Sr. Staff Solutions
Architect, Databricks

Ajay Gollapalli

Director, Ecosystems,
Informatica

AGENDA
Build your Analytics in DB SQL with Trusted High-Quality Data Using IDMC

- Informatica & Databricks Partnership

- How Informatica Adds Value On Databricks

- Product integration roadmap

- Demo - Informatica Cloud Data Governance with Databricks Unity Catalog

- Q&A

Informatica & Databricks Partnership

Informatica extends the value of your Databricks investment
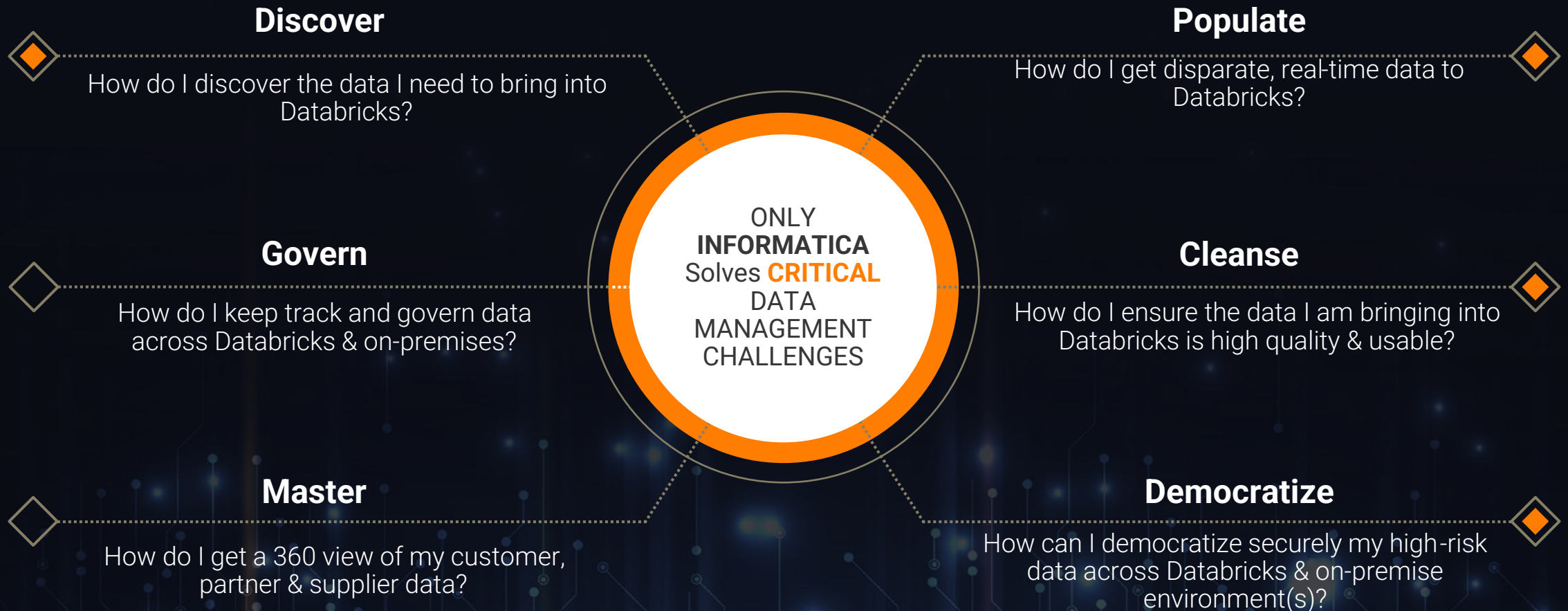

160+ Joint Customers

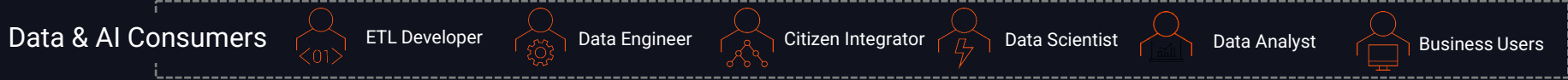**Data Integration Partner of the year 2024**

Top 3 fastest growing on Databricks

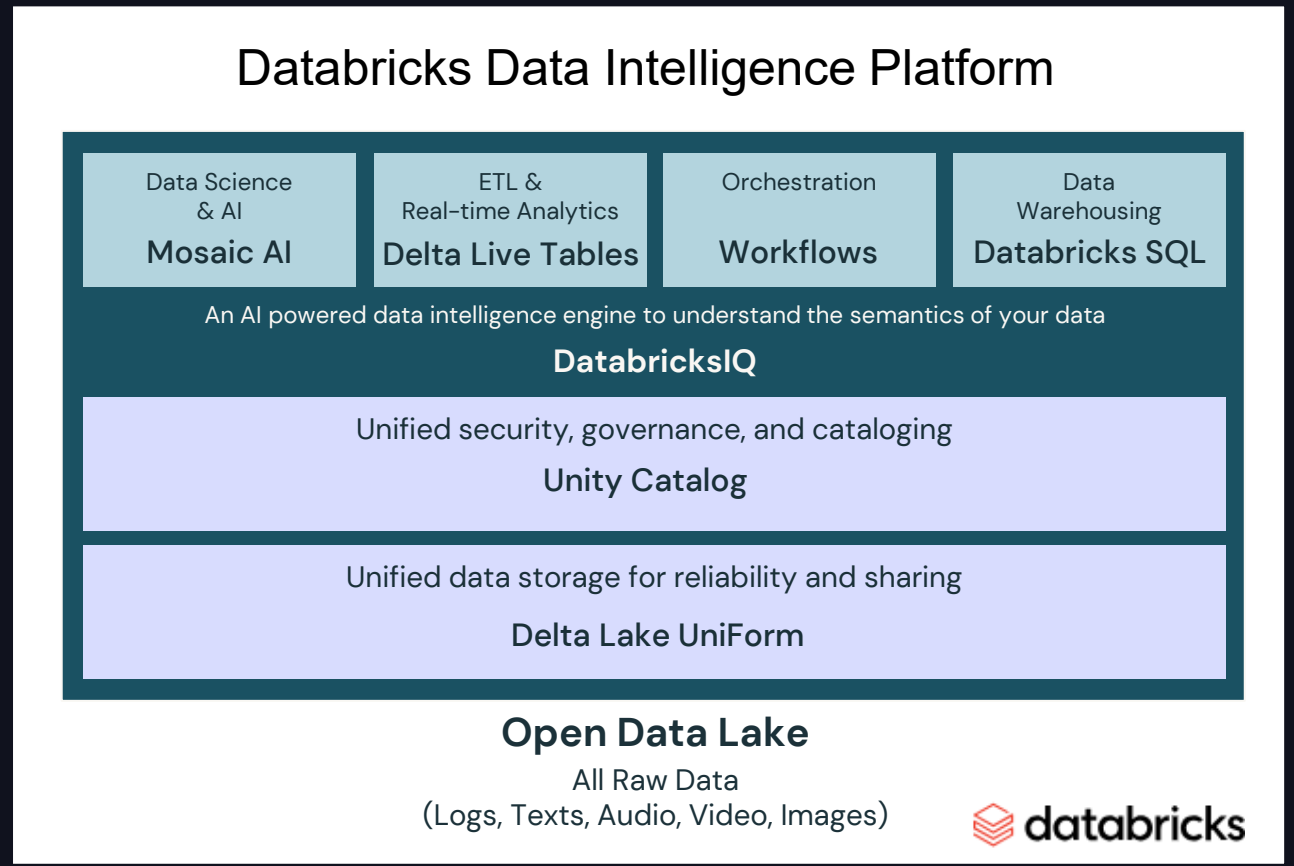# How Informatica Adds Value On Databricks

Informatica Solves Critical Data Challenges for Databricks Customers

Discover

How do I discover the data I need to bring into Databricks?

Populate

How do I get disparate, real-time data to Databricks?

Govern

How do I keep track and govern data across Databricks & on-premises?

ONLY **INFORMATICA** Solves **CRITICAL** DATA MANAGEMENT CHALLENGES

Cleanse

How do I ensure the data I am bringing into Databricks is high quality & usable?

Master

How do I get a 360 view of my customer, partner & supplier data?

Democratize

How can I democratize securely my high-risk data across Databricks & on-premise environment(s)?

Informatica

Informatica Integrates with Databricks

Data & AI Consumers — ETL Developer · Data Engineer · Citizen Integrator · Data Scientist · Data Analyst · Business Users

Informatica — Intelligent Data Management Cloud

- MDM & 360 APPLICATIONS
- DATA INTEGRATION
- DATA QUALITY & PROFILING
- API & APP INTEGRATION
- DATA CATALOG
- CLOUD DATA MARKETPLACE
- GOVERNANCE & PRIVACY

Databricks Data Intelligence Platform

Data Science & AI	ETL & Real-time Analytics	Orchestration	Data Warehousing
Mosaic AI	Delta Live Tables	Workflows	Databricks SQL

An AI powered data intelligence engine to understand the semantics of your data
DatabricksIQ

Unified security, governance, and cataloging
Unity Catalog

Unified data storage for reliability and sharing
Delta Lake UniForm

Open Data Lake
All Raw Data
(Logs, Texts, Audio, Video, Images)

databricks

Informatica

Product Integration Roadmap

Informatica + Databricks – Integration Roadmap

Now Available on Partner Connect

Generally Available for Customers

Cloud Data Governance & Catalog

- Support for Scanning Unity Catalog data assets

- Integrated with Unity Catalog API

- Integrated with Partner connect (Announcing Today)

Cloud Data Marketplace

- Integrated Shopping cart experience for Databricks assets to Federated users

- Support for Databricks SQL endpoint for Copy into

- Support for Delta Lake & SQL Unity Catalog as endpoint

Upcoming Features

- AI Governance Scanning for Catalog

- Databricks Volumes support

- Databricks Vector Index Scanning

- Databricks Delta share support

Data Governance – The Baseline View
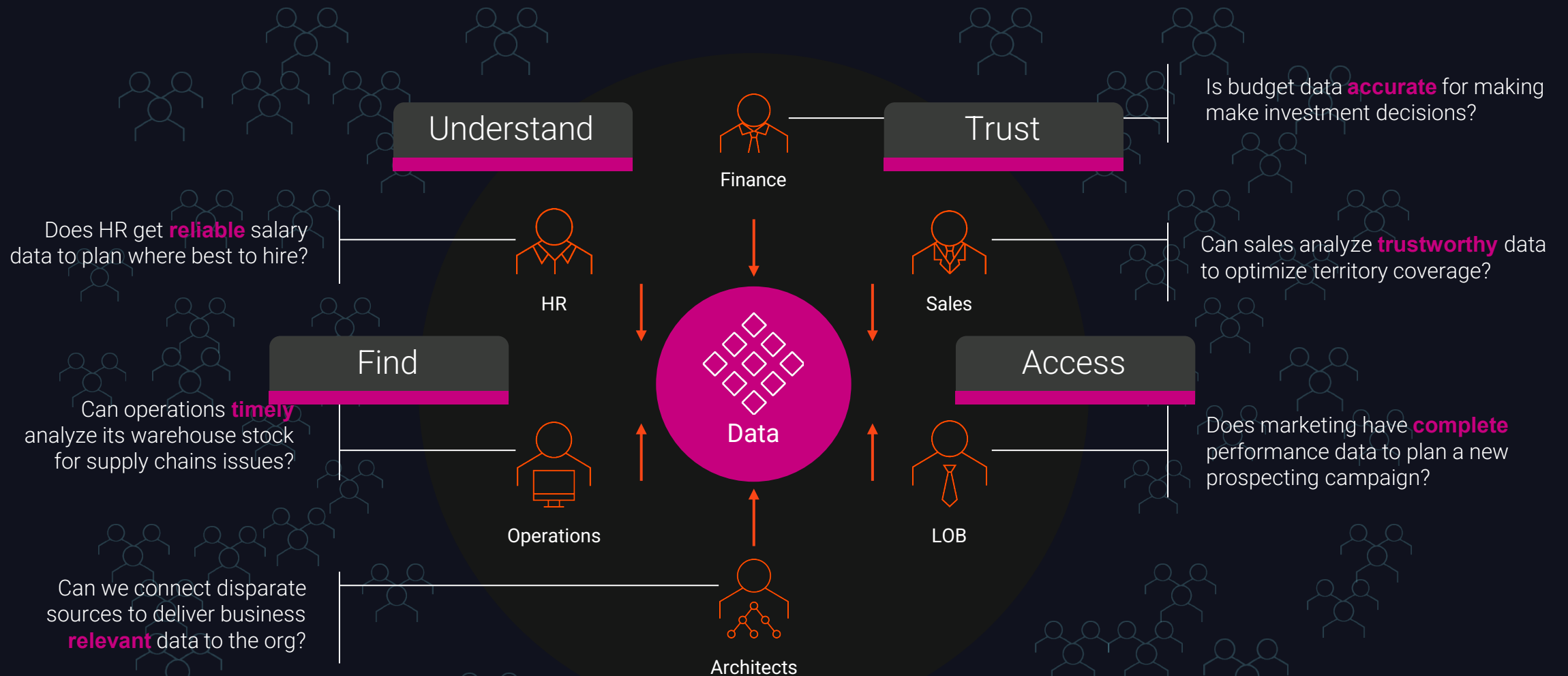
A Focus on Risk and Compliance

DATA+AI SUMMIT

Governance for Risk and Compliance

Making data safe with auditing and reporting for transparency

Risk & Compliance

Manage risk and comply with
regulatory mandates and policies

Requirements for Data Democratization

Is budget data **accurate** for making make investment decisions?

Does HR get **reliable** salary data to plan where best to hire?

Can sales analyze **trustworthy** data to optimize territory coverage?

Can operations **timely** analyze its warehouse stock for supply chains issues?

Does marketing have **complete** performance data to plan a new prospecting campaign?

Can we connect disparate sources to deliver business **relevant** data to the org?

Data Analytics Governance With Data Sharing

Ensuring data availability, reliability and governance

Risk & Compliance

Manage risk and comply with regulatory mandates and policies

Analytics & Data Sharing

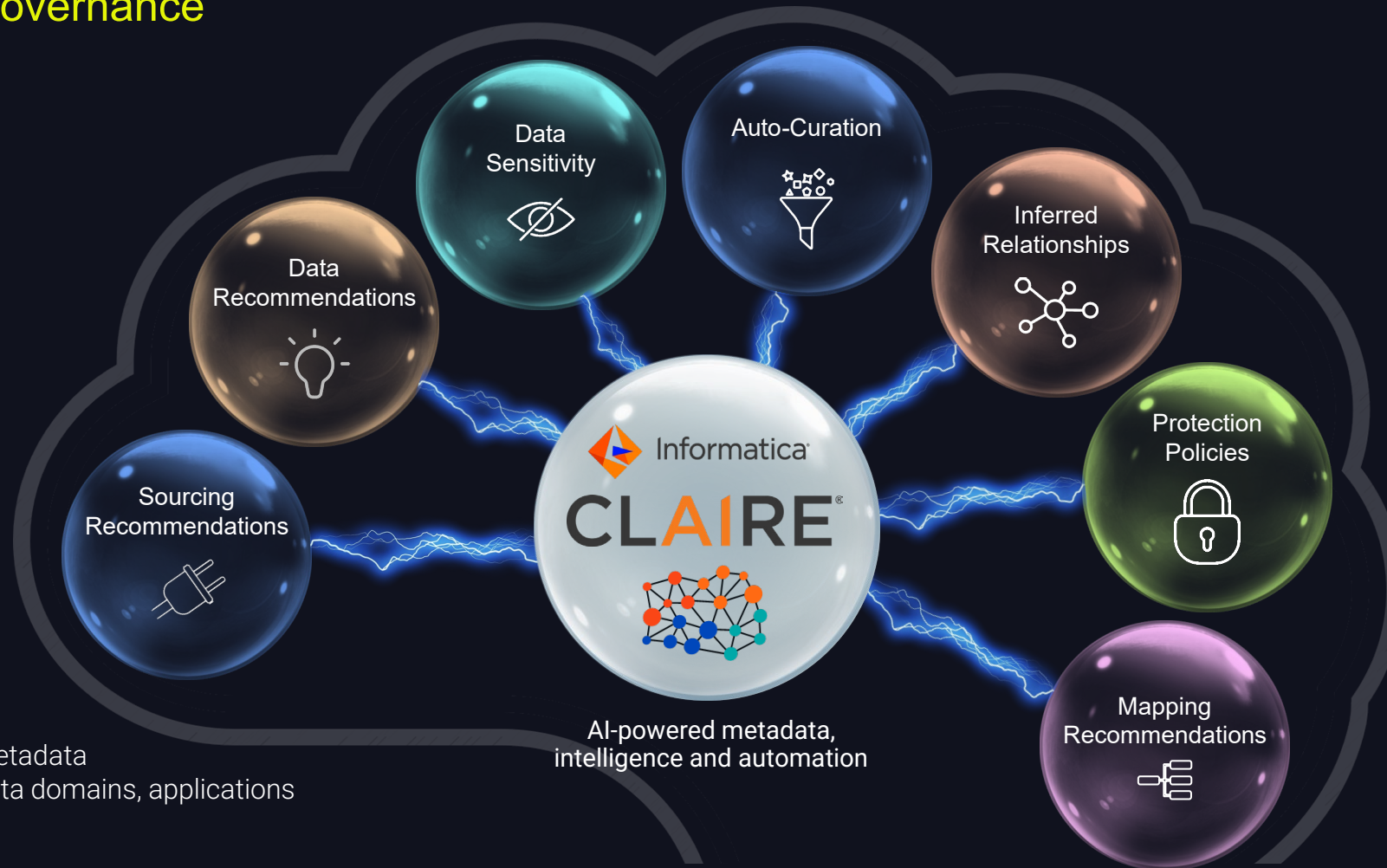Open data for business with actionable insights and automation

What's next to Modernize Data Governance?

Automate Modern Data Governance needs at Scale

Operationalize trust

Persona-aligned views on health

Automation for speed and scale
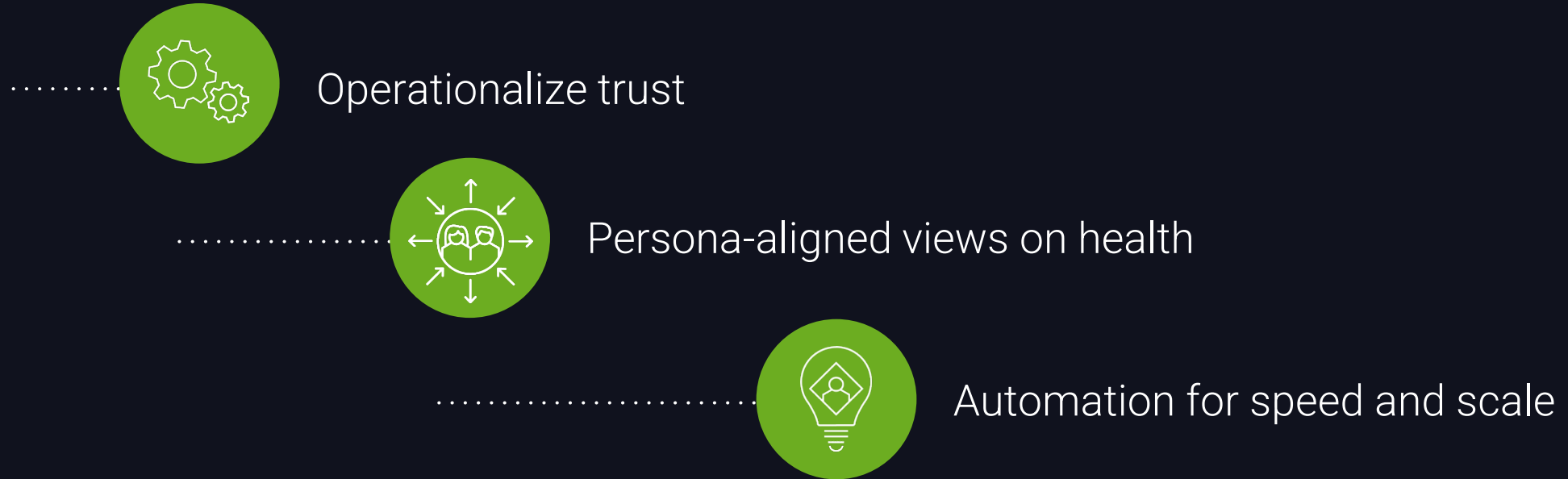
Manage and scale high-quality data pipelines for Analytics & AI

The Need for an Operational View on Trust

Business

Monitor how data is consumed, protected and compliant with policies and regulations

Data

Provide visibility into data health — identify impact of any issue for preventative or remedial action

Pipelines

Optimize availability, performance and capacity — cost-effectively and efficiently — on any platform

Modern Data Governance

Ensure Data Availability, Reliability & Governance

Risk & Compliance

Manage risk and comply with regulatory mandates and policies

Analytics & Data Sharing

Open data for business with actionable insights and automation

Intelligent Data Observability

Optimize availability, performance, capacity — cost-effectively and efficiently

CLAIRE® AI is the Brainpower Behind

Modern Data Governance

Data Sensitivity

Auto-Curation

Data Recommendations

Inferred Relationships

Sourcing Recommendations

Informatica®
CLAIRE®

AI-powered metadata, intelligence and automation

Protection Policies

Mapping Recommendations

40 petabytes of active metadata
Organization, industry, data domains, applications
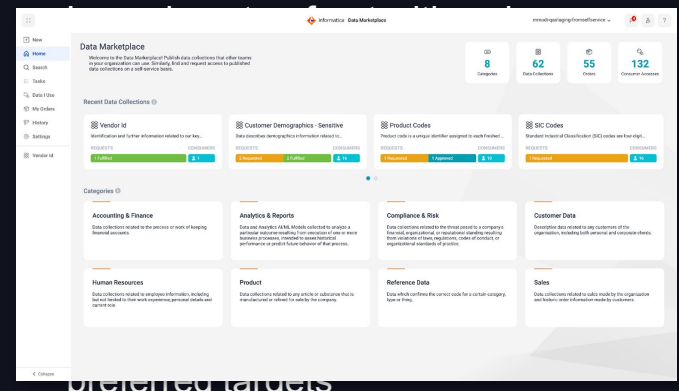
Data Governance and Catalog
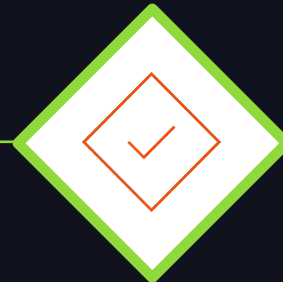### Understand and Govern
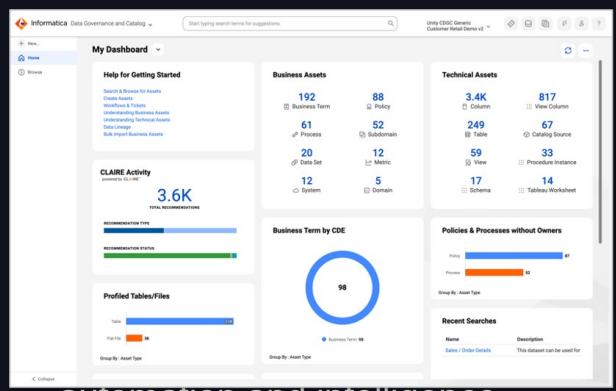
automation and intelligence

Data Marketplace
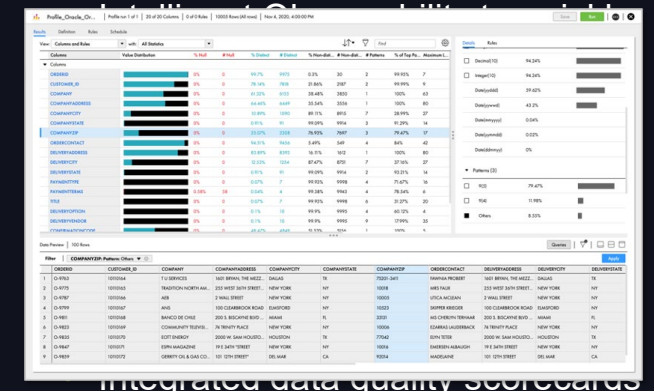### Share and Democratize

preferred targets
- Track fulfillment with full visibility

Data Quality and Observability
Cleanse and Trust

integrated data quality scorecards
for trust-based governance

Demo

Empowering all Data Consumers & Data Owners

Automate Modern Data Governance at Scale

Governance

Simplify trusted data delivery with a unified set of cloud-native services to scale on demand, as needed

Democratization

Accelerate time to value for data-driven outcomes, connecting data producers and consumers

Observability

Achieve reliable analytics insights by monitoring pipelines and improving data accuracy

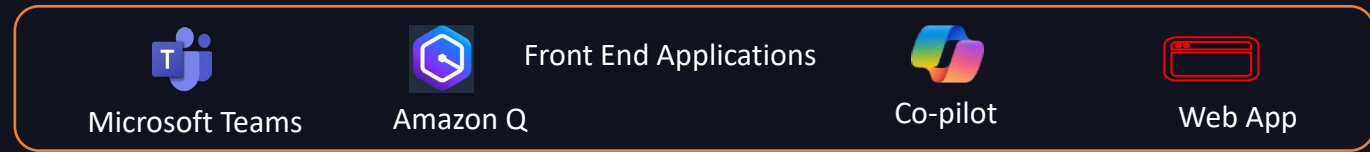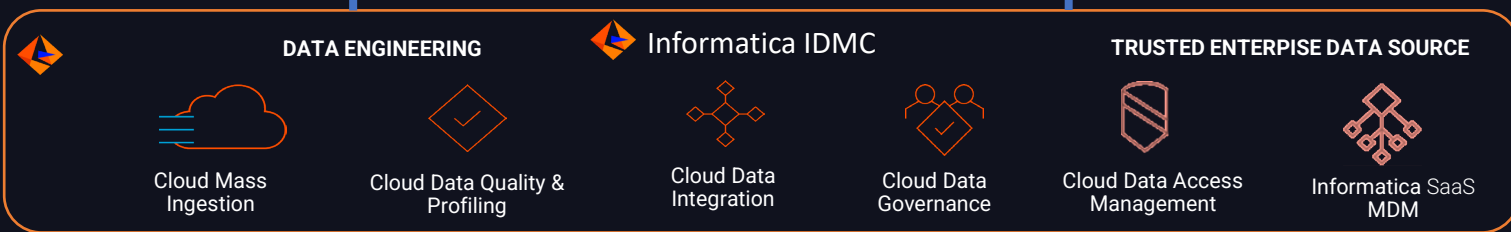Informatica

CDGC & Unity Catalog Customer Value

CDGC + Unity Catalog

Unified Metadata
Management

Consistent Business
Glossary

Data Lineage and
Impact Analysis

Policy Enforcement
and Compliance

Customer
Benefits

Enhanced Data
Quality Management

Facilitates Cross-
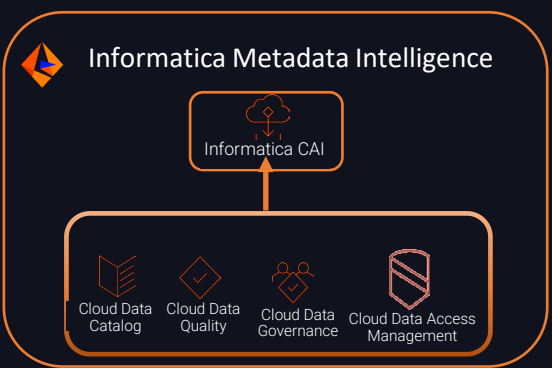Functional Collaboration

Holistic View of Data
Ecosystem

Efficient Change
Management

Informatica Integrates with Databricks DBRX

Front End Applications
Microsoft Teams · Amazon Q · Co-pilot · Web App

User Query → Response

RAG CHAIN
Informatica CAI | Prompt Deconstructor (DBRX) → Informatica Metadata Intelligence API → Data Fetch Agents → External API (REST / SOAP) → Response Constructor (DBRX)

Planner & Orchestrator — Executors / Agents

Data Sources
Sales, Service, Marketing — salesforce, Adobe, Adobe Experience Manager, Magento, Marketo, servicenow, ORACLE CRM ON DEMAND, eloqua, ORACLE NETSUITE, zendesk, ZOHO, Twitter, Facebook, LinkedIn

Apps · Equipment · Cloud
Log files · Mobile · Machine Data
Application Servers · Mainframe · tabases

Informatica Metadata Intelligence
Informatica CAI
Cloud Data Catalog · Cloud Data Quality · Cloud Data Governance · Cloud Data Access Management

Databricks Data Intelligence Platform
Databricks SQL → Vector Index → Vector Search Endpoint ↔ Databricks DBRX & Model Serving

DELTA LAKE

DATA ENGINEERING — Informatica IDMC — **TRUSTED ENTERPISE DATA SOURCE**
Cloud Mass Ingestion · Cloud Data Quality & Profiling · Cloud Data Integration · Cloud Data Governance · Cloud Data Access Management · Informatica SaaS MDM

22
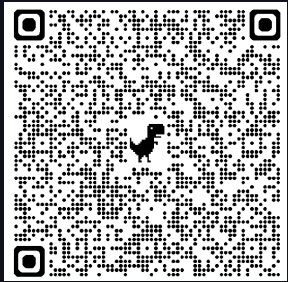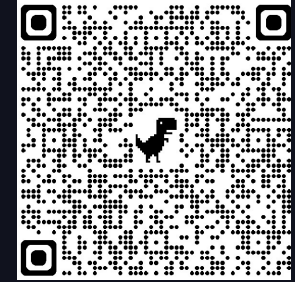
DATA⁺AI SUMMIT

Scan QR codes to learn more

Cloud Data Integration Free Service via Partner Connect

Learn more on Informatica with Databricks

Read about our Latest Announcements

Thanks

DATA+AI SUMMIT