Democratizing and Monetizing Data at M Science with the Data Intelligence Platform
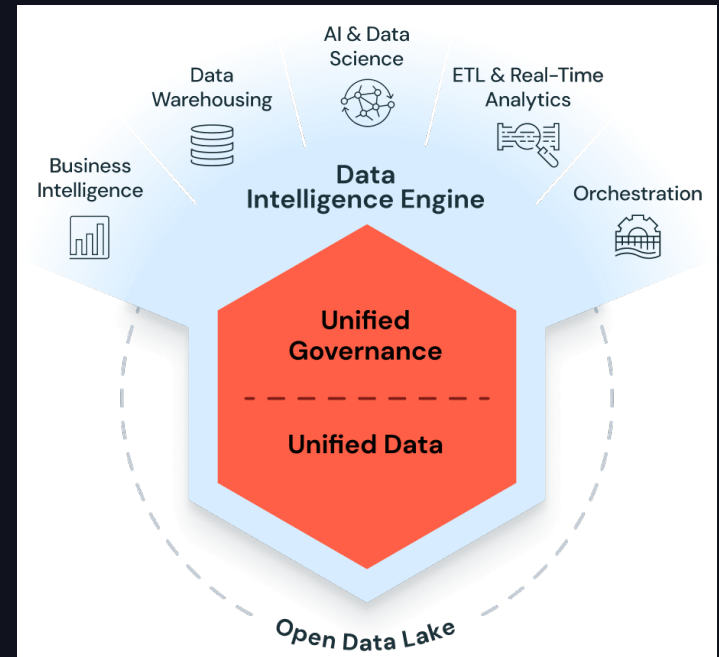
Data Intelligence Platform. What is it?

Data Intelligence Platform?

What is it?

Databricks combines generative AI with the unification benefits of a lakehouse to power a Data Intelligence Engine that understands the unique semantics of your data.

What does M Science Do?

What does M Science do?

M Science Transforms Data Into Actionable Insights

providing

Data-Driven Research to Empower Decisions

DATA AI SUMMIT

What does M Science do?

With a bit more color...

Over 200 Data Sources

- Transaction Data
- Receipt Data
- Web Traffic Data
- Hundreds of Web crawlers
- Secret Data

Intelligent Data Platform

- Specialized Analyst Teams
- Databricks Lakehouse
- >20 years of experience

Actionable Insights

- In depth research reports
- Analyst Curated Dashboards
- Curated Analyst Data Feeds
- Curated Raw Data

What does M Science do?

Monetize your de-identified anonymous data...

Data monetization brings high-margin revenue (85%+) in a short period of time
PaymentsSource

Frame a winning strategy: Effective data monetization strategies recognize data as an evolving asset living and growing within a living enterprise — NOT as a static asset that depreciates in value
KPMG

New data-driven value streams start to materialize once a company's business leaders realize that one company's trash is another company's treasure
CardinalHealth

Every company operating today is a data company — most are leaving money on the table
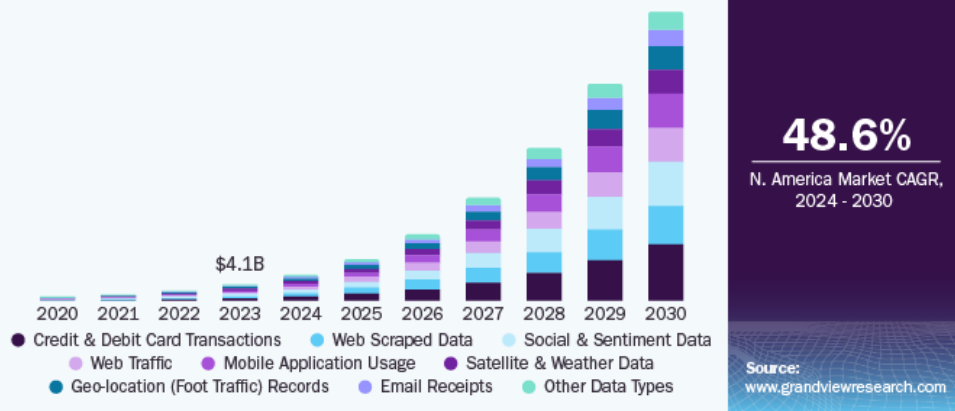**MITSloan** Management Review

It is expected that 163 zettabytes of data will be generated globally by 2025
Deloitte.

Data monetization and individual privacy can coexist with a comprehensive compliance framework

North America Alternative Data Market
Size, by Data Type, 2020 - 2030 (USD Billion)

$4.1B

2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030

- Credit & Debit Card Transactions
- Web Scraped Data
- Social & Sentiment Data
- Web Traffic
- Mobile Application Usage
- Satellite & Weather Data
- Geo-location (Foot Traffic) Records
- Email Receipts
- Other Data Types

GRAND VIEW RESEARCH

48.6%
N. America Market CAGR, 2024 - 2030

Source: www.grandviewresearch.com

Why the Data Intelligence Platform?

Why the Data Intelligence Platform?

Going back to the beginning...

Greenplum

- Noisy Neighbor Syndrome

- Expensive On Prem CapEx

- Transactional scale

Why the Data Intelligence Platform?

The Past...

Data Lake

In 2016 Databricks on S3

- By 2019 >60,000 tables

- Limited Data Catalog Capabilities

- Orchestrated, Self-managed Spark

- Notebooks for Data Engineering

- Millions of data files in S3, not tagged

Why the Data Intelligence Platform?

The Recent Past...

Data Lake

Databricks on S3

- >100,000 tables

- Orchestrated, Self-managed Spark

- Python Library for ETL

- Millions of data files in S3, not tagged

Orchestration

Airflow from Astronomer

- Focus on DAGs

- Reasonable costs

- Use, don't run Airflow

Data Warehouse

Snowing and chilly

- Well organized

- Expensive at scale

- Live connection for Tableau

- Source of Truth for Share

The Data Intelligence Platform!

The Data Intelligence Platform?

Medallion Architecture

Bronze

- Ingestion
  - S3
  - Delta Share
  - sFTP
  - Snowflake Share
  - ...
- Monitored Wild West

Silver

- Product Pipelines
  - Tagged & Normalized
  - Shared Python ETL Library

- Standardized and Monitored

Gold

- Curated Data
  - Genie Spaces
  - Delta Share

- Analyst Playground

- Marketplace

The Data Intelligence Platform?

Unity Catalog

Databricks S3 Lakehouse

- The world is your oyster...

    ...

- Federated

    ...

- Catalog and manage EVERYTHING

    ...

- Trusted data

The Data Intelligence Platform?

Serverless SQL, Jobs & Liquid Clustering, Monitoring out of the box

Remember Unity Catalog?

- Serverless SQL and Jobs

    …

- Delta Share

    …

- Liquid Clustering & UniForm

    …

- People are expensive
    - Platform management is undifferentiated

- Lakehouse Monitoring
    - Out of the box differentiated win with NO people

The Data Intelligence Platform?

Databricks AI for Productivity

AI Assistant

- Text to SQL

    ...

- Context Aware

    ...

- Productivity Rocket

    ...

Genie Spaces

- AI for Business Questions

    - Demo...

DATA AI SUMMIT

M Science: RAP

The Data Intelligence Platform?

M Science RAP

Research, Author & Publish

- Long form Research with AI

  …

- Demo

DATA·AI SUMMIT