

# How AI-Driven Master Data Management Can Speed Up Your Business Outcomes

Costa Mikhael – Informatica MDM Leader - APAC

October 2024

Where data  
& AI come to 

# Is your business ready for the next revolution



## Data for All

Allow everyone in the organization to be data driven, regardless of technical ability



## Create Usable Data

Enable decentralized data producers to create, document, catalog, and test new data assets



## Drive Exponential Productivity Growth for Your Data Team

Drive exponential growth in productivity of your data professionals with  
**Artificial Intelligence (AI)**

# Flawed and Inconsistent Data Slows Down Business

There Exists a Fragmented Landscape of Applications, Data, People

## Applications/SaaS



ERP



SaaS



IOT



CRM

~500+

cloud apps in use in most organizations

## Data



Structured Data



Real-time



Unstructured Data



Batch Data

10 petabytes

of data stored by average enterprise.  
180 zettabytes total by 2025

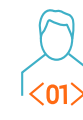
## People



Data Architect



Business User



Developer



Business Leader



CIO/IT Leader



CDO



Data Scientist

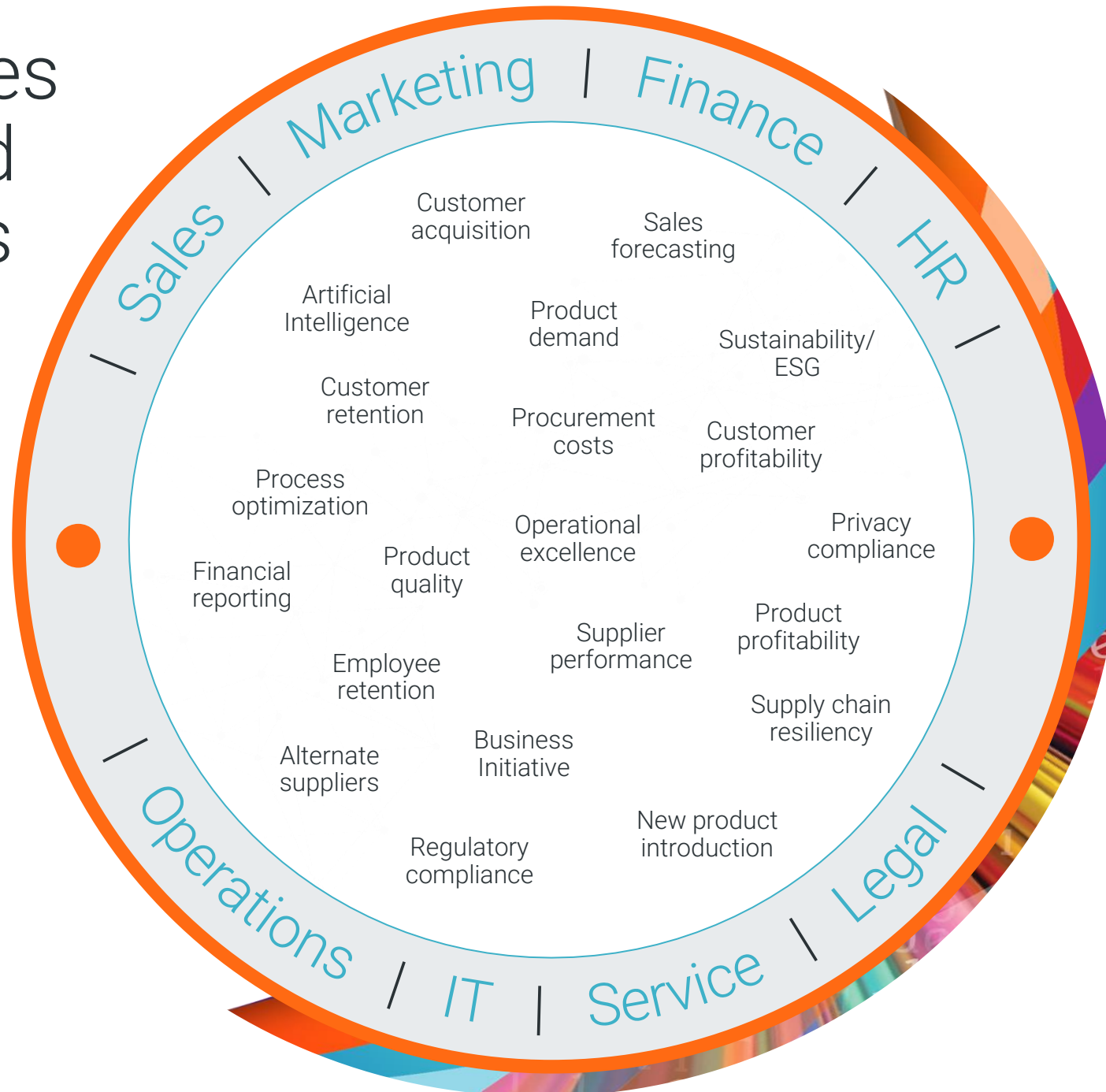


Data Engineer

88% of execs

believe technology democratization is critical to ignite innovation

...and Creates  
Fragmented  
Experiences

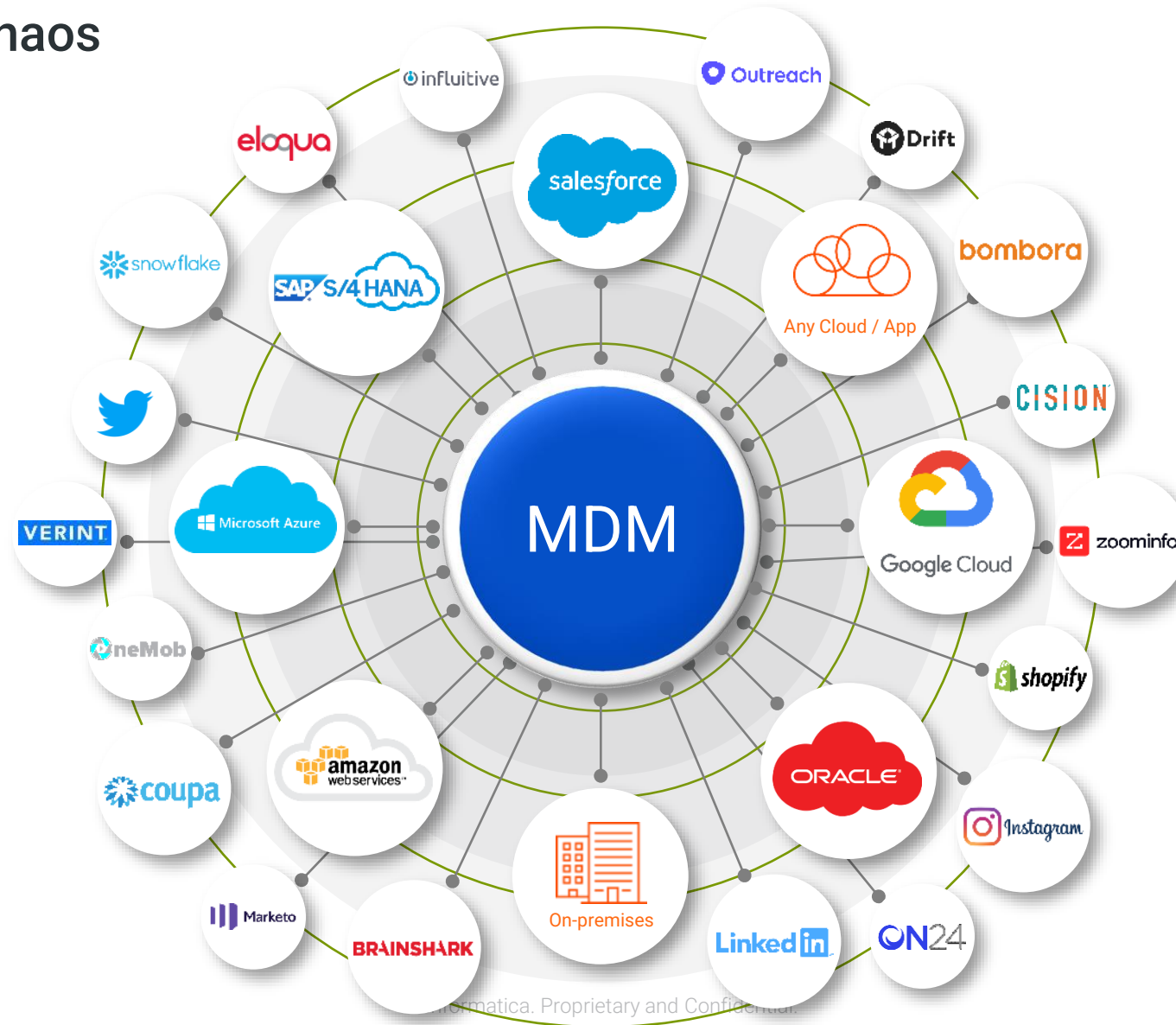




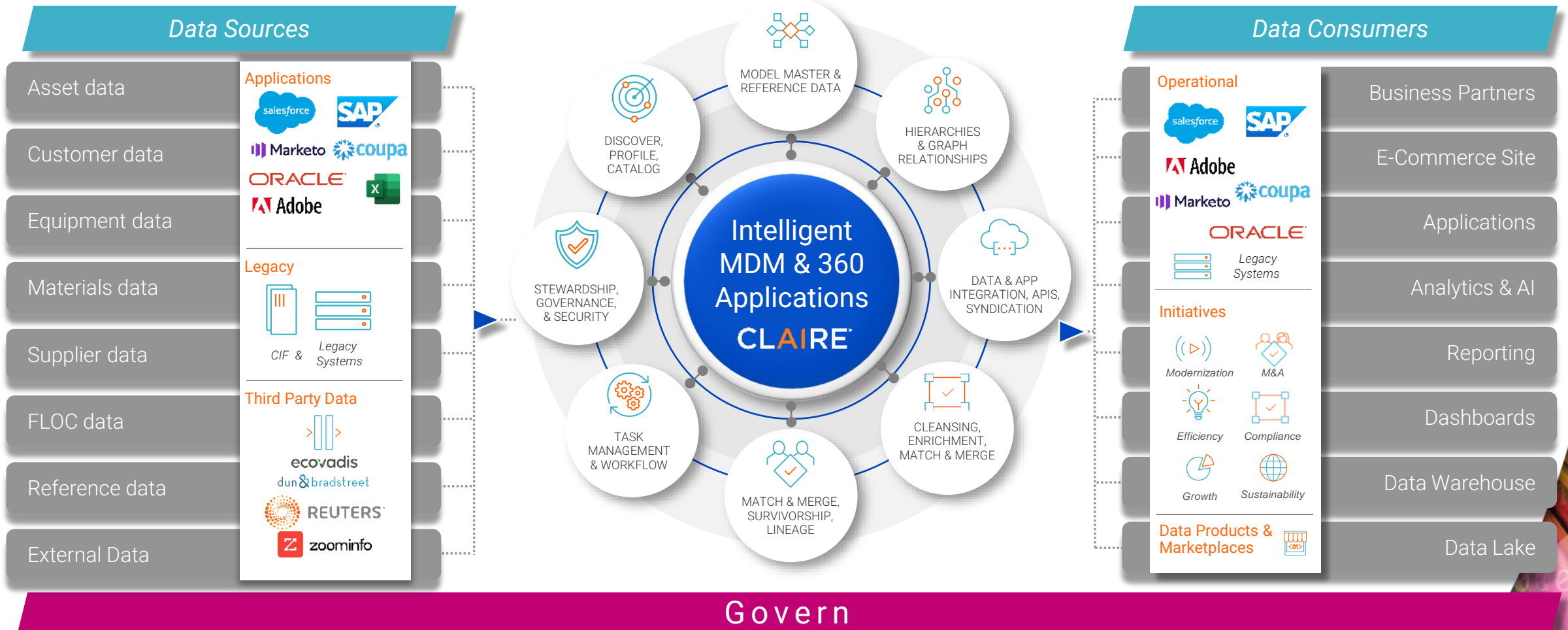


# Master Data Management

Bringing order to chaos



# Multi-Domain Master Data Management



# Intelligent MDM & 360 Applications

  
ESG Extension for  
Supplier 360

  
Product 360 for  
Life Sciences


  
Customer 360 for  
Life Sciences


  
Customer 360  
for Healthcare


  
Customer 360 for  
Insurance


  
MDM 360 for  
Retail


  
Legal Entity 360  
for Banking


  
Customer 360

  
Product 360


  
Supplier 360

  
Reference 360

  
Finance 360

  
Location 360

  
Material 360

  
Any Domain

## All-in-One MDM Services



Ingress



Modeling



Hierarchy  
Management



Data  
Quality



Enrichment



Matching



Survivorship



Stewardship &  
Governance



Search &  
Reports



Workflows



Task  
Management



Security



Egress

## SHARED IDMC MICROSERVICES

Cloud Data  
Integration

Cloud Data  
Quality

Cloud  
Application  
Integration

Cloud Data  
Catalog

Identity  
Service

Session  
Service

Federated  
Repository  
Service

Secrets  
Management

Configuration  
Management

# CLAIRE<sup>®</sup>

## Connectivity

10K+ METADATA-AWARE CONNECTORS



# CLAIRE<sup>®</sup> AI within MDM

## Intelligent MDM & 360 Applications

Every Master Data Domain

### All-in-One MDM Services

← SHARED IDMC MICROSERVICES →

Cloud Data  
Integration

Cloud Data  
Quality

Cloud  
Application  
Integration

Cloud Data  
Catalog

Identity  
Service

Session  
Service

Federated  
Repository  
Service

Secrets  
Management

Configuration  
Management

**CLAIRE<sup>™</sup>**

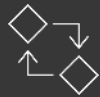
AI-Powered Metadata Intelligence & Automation



Identity  
Matching



Product  
Matching



Schema  
Mapping



Rule  
Recommendation



Dataset  
Recommendations



Product  
Classification



Data Set  
Similarity



Glossary  
Association



Entity  
Extraction



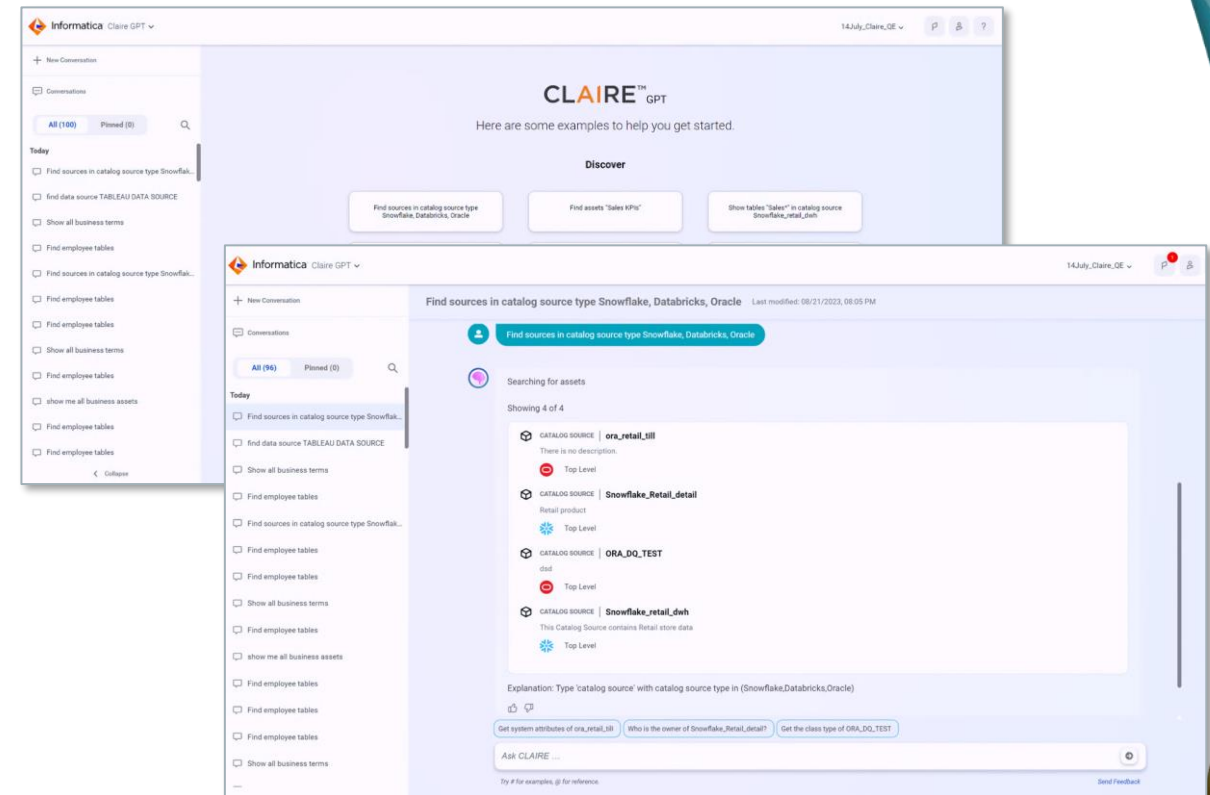
Data Domain  
Inference

### Connectivity

Connect to All Systems Containing or Needed Master Data

# CLAIRE GPT – It's time

- **Data Discovery:** Use natural language queries to easily find relevant data assets, tables, columns, etc., including data discovery with specific characteristics like PII/sensitive data
- **Metadata Exploration:** Search and explore business terms, definitions, data lineage, data quality scores, data owners/stakeholders associated with data assets.
- **Data Exploration:** Explore data assets with natural language questions to understand trends, KPIs, top customers, etc. without needing SQL knowledge.
- **ELT:** Generate new data products by creating ELT pipelines using natural language on data warehouses like Snowflake, AWS Redshift, Azure Synapse, Google Big Query, Databricks Delta Lake and more.
- **Product Help:** Answer product and PAM related questions
- **Self-Service for Business Users:** Enable non-technical business users to find and understand data on their own by abstracting away the need for SQL.
- **Informatica LLMs:** CLAIRE GPT is powered by fine-tuned open-source large language models created for data management.



# There is NO AI Without Trusted Data

## Why AI Fails?

### Poor Quality Data

- Data is critical when it comes to securing successful AI outcomes
- Trusted data helps AI learn to make a recommendation or perform a task
- The more fit for use data sourced, the better the outcomes

### Do you want to prevent AI failure?

#### Cost of Poor Data Management<sup>1</sup>

- *"Average cost of poor data quality at \$12.8 million per year"*
- *"AI models depends entirely on the accessibility, quality, granularity, and latency of your organisation's data"*

1. "The steep cost of a poor data management strategy." (2023), [CIO.com](#)

2. "Famous Analytics and AI Disasters." (2023), [CIO.com](#)

#### Data and AI Bias

- 86% of business and IT decision-makers believe data bias will become a bigger concern as AI/ML use increases
- Yet over a third (34%) have not looked at technology and tools to help tackle it

#### MS Chatbot Spew Racist Tweets<sup>2</sup>

- Used uncleansed Twitter interactions for training ML models
- Within 16 hours, more than 95,000 tweets posted, and rapidly turned overtly racist, misogynist, and anti-Semitic

#### Bad AI Leads to Attorney Sanctions<sup>2</sup>

- Used generative AI to find prior cases to support a case
- At least 6 of the AI findings did NOT exist
- As of June 2023, Schwartz was facing possible sanctions by the court

#### Misdiagnosis Due to Errant AI<sup>2</sup>

- Hospitals used AI to diagnose & triage COVID-19 patients faster
- AI model was flawed because it was trained on the wrong data set
- Mislabeled data or data from unknown sources was a common culprit



**Reduced** cost improve operation efficiency for managing locations and facilities

**Improved** student experience with enterprise of view of student information

**Improved** Insight and reporting



**Protecting Children** – Collaborate with other services and professionals

**Identify** Needs, issues, and vulnerability

**Make Informed Decision** about safety and well being of children



New Zealand Government

**Standardize** on National procurement catalogue for medical devices/equipment

**Better Access:** Health District Boards

**Accurate:** Spend reporting





Reduced IT cost by  
**25-30%**

**50% reduction**  
of duplicate  
customer records  
leading to improved  
response rates,  
segmentation and  
personalized  
communications

Customer loan  
approvals  
provided in  
**3 minutes**  
versus 4 to 6  
weeks

Connect & reconcile  
**50 million customer  
records**  
to reach customers in  
a more personalized &  
engaging manner

**12x higher  
revenue**  
and identified  
more cross-sell  
and up-sell  
opportunities

Improved KYC  
data quality  
**35% to 100%**

Created an  
additional  
**20,000 hours**  
without  
adding new  
FTE

**15%-20%**  
reduction in  
support tickets  
related to user  
issues

**5x faster  
coverage**  
and suspicious  
transaction  
reporting for  
AML  
compliance

**38% reduction**  
in time to  
create and  
validate  
company  
record

**+19.6%  
revenue  
premium** from  
loyalty  
customers

**2% increase** in  
customer  
satisfaction  
(CSAT)

**\$0.5M-\$1M  
saved**  
managing and  
remediating  
data issues

**14 Days**  
from start to  
create a  
Trusted  
Customer  
Profile