Modern Data Governance:
How to Improve Business Results With Automation and AI

Nate Rain
Data Governance Specialist

Donna Heimbaugh
Senior Product Specialist

Where data & AI come to LIFE

From Digital Transformation to Scaling Digital Business

Operating a digital business
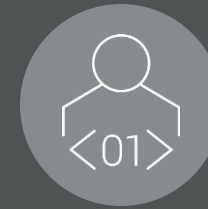
48%

of enterprises are a digital business[1]

From tech suppliers to trusted ecosystems

79%

report trust programs are a priority[1]

Prioritizing data governance

52%

say improving data governance is the #1 priority for CDOs[2]

1. IDC, "The Five Levers Required to Scale a Digital Business," Doc #DR2023_GS2_MW, March 2023
2. Informatica, "CDO Insights 2023: How to Empower Data-Led Business Resiliency," December 2022

Informatica

Common Challenges in Data Governance

Where is the data coming from?

Data Engineer

Data Architect

Who can provide feedback?

Can I trust this report?

Executives

Data Governance

How do we get access to data?

Which definition is correct?

Stakeholder

Where can I find PII data?

Scientist

Analyst

Who owns this data?

Data governance is about enabling organizations to **DEFINE, DISCOVER, MEASURE, COLLABORATE** on, and **SHARE** data concepts.

Informatica

3

Modern Data Governance

Ensure data availability, reliability and governance

Risk & Compliance

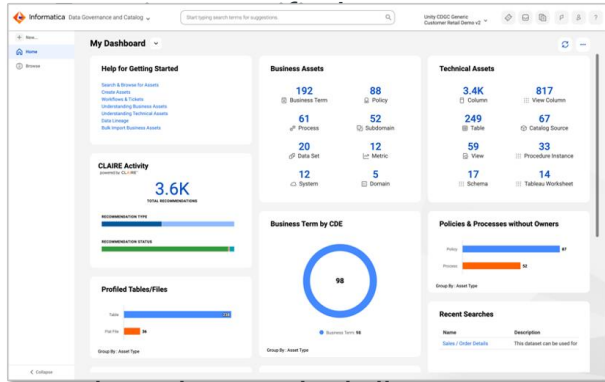Manage risk and comply with regulatory mandates and policies

Analytics & Data Sharing

Open data for business with actionable insights and automation

Intelligent Data Observability

Optimize availability, performance, capacity – cost-effectively and efficiently
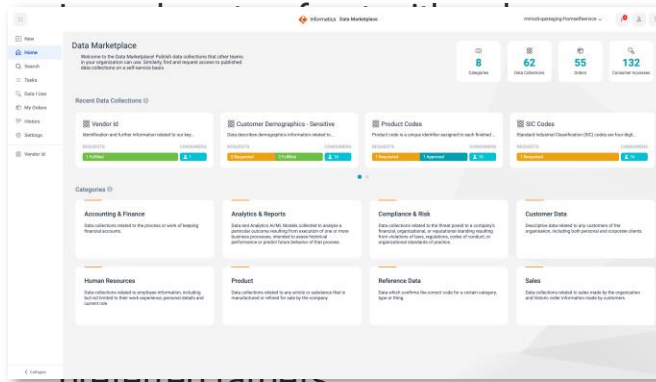
Data Governance and Catalog
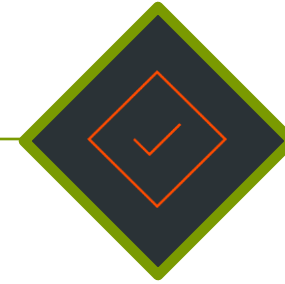

automation and intelligence
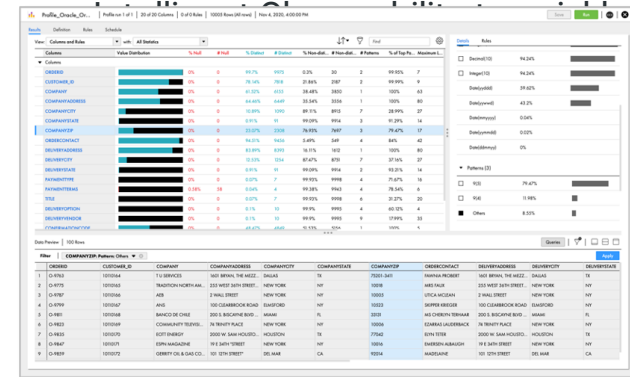
Data Marketplace

preferred targets
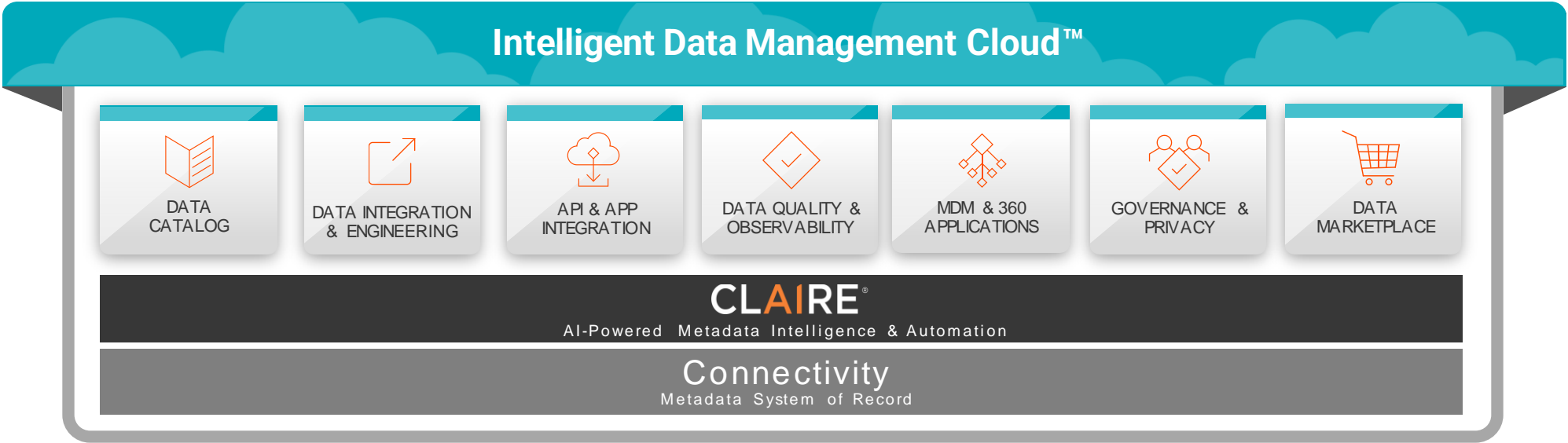- Track fulfillment with full visibility

Data Quality and Observability

integrated data quality scorecards for trust-based governance

The Industry's Only Intelligent Data Management Cloud

Intelligent Data Management Cloud™

| DATA CATALOG | DATA INTEGRATION & ENGINEERING | API & APP INTEGRATION | DATA QUALITY & OBSERVABILITY | MDM & 360 APPLICATIONS | GOVERNANCE & PRIVACY | DATA MARKETPLACE |

CLAIRE®
AI-Powered Metadata Intelligence & Automation

Connectivity
Metadata System of Record

Multi-Cloud

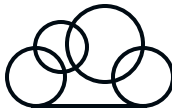aws Microsoft Google Cloud ORACLE Cloud

Hybrid

On-premises Enterprise Cloud Serverless

Informatica®

Office of the CDO Use Case Investments

Data intelligence to enable data-driven transformation

BUSINESS IMPERATIVES

| GROW THE BUSINESS AND INCREASE AGILITY | IMPROVE CUSTOMER EXPERIENCE | REDUCE COSTS | RISK AND COMPLIANCE |

USE CASE INVESTMENTS

| Meet the needs of the business by **aligning data strategy with business priorities** | Democratize data by empowering consumers to **find, understand, trust and access relevant data** | Reduce costs and grow revenues with **self-service access to trusted data** for business insights | Ensure **data is trusted and used responsibly** enforcing governance policies and standards |

The Need for Automation and AI

Data professionals spend around **38%** of their time on data preparation and cleansing, rather than focusing on higher-value work including selecting models, interpreting results and generating and presenting meaningful insights.[1]

1. Anaconda, "2022 State of Data Science Report", September 2022

Informatica

Data Governance

Business-focused Information Consumption and Management

- **Intuitive Relationship Visualizations of governance artifacts and the associated business impact to effectively manage and operate governance programs**
- Multi-Language for Global rollouts with support for 6 languages
- Customizable and Personalized Dashboards
- Multi-hop graph-based searches to browse and navigate complex relationships

Collaboration

- **Crowd-sourced knowledge for continuous feedback, ratings, certifications**
- **Governance and Technical Asset Ratings to increase trust and confidence**
- Discussion Threads for informal interactions/clarifications
- Workflows and Ticket management for structured interactions

Automation and Integrations to accelerate business insights

- **Quality overlays for Business and Technical Lineage to improve understanding and accelerate Quality Coverage**
- Data Quality Self Service using NLP-based Rule generation

Data Access Management (Privitar)

- **Metadata-driven policies with context**
- **Comprehensive controls to enforce policy**
- **Universal access policy enforced across hybrid IT**
- **Audit and reporting for compliance posture**

CLAIRE ®

- Intelligent Glossary Associations to effectively enrich scanned assets with business context
- **Centralized dashboard to review and manage CLAIRE generated Glossary Associations and Classifications**
- Data Quality Automation to assess sources against enterprise standards

Up to **80%** reduction in support requirements by eliminating cross-dependencies

1,000s of hours of manual work saved

Up to **70%** reduction in time to find data

Data Governance (Catalog)

Advanced Scanners and Curation

- Broadest metadata coverage across all key ecosystems: Snowflake, Microsoft, AWS, Google, SAP, Oracle and Salesforce/Tableau and more
- **Advanced Scanning capabilities** include extracting data lineage from Stored Procedure Scanning, Python Code Scanning, and Embedded SQL in Data Pipelines
- Bring your own metadata with extendable Custom Scanner Framework
- **Bulk Curation to eliminate human task and accelerate data labeling and glossary association outcomes**
- **High-granularity Sensitive Data Classification for visibility, tracking and reporting**

CLAIRE®

- 225+ Pre-built based Data Elements like national IDs, dates, locations, names
- Classification builder to create custom classifications
- **CLAIRE generated classifications**
- **Aggregate elements into entities to classify and discover business entities**

Lineage

- **Automated Column-Level Lineage. views with business overlays including Data Quality, Policies, Glossaries, and Processes**
- Advanced lineage scanning including parsing transformations hidden in SQL code, ETL/ELT code, Python Code, BI Tools and more
- **Detailed Impact Analysis reports for both upstream and downstream changes**

Up to **400** auto-generated classifications in minutes

Up to **80%** reduction in metadata extraction time

Up to **100s** of hours of manual work saved

Cloud Data Marketplace

Data Product Lifecycle Management

- Distributed Data Asset Onboarding – Any catalog/any data set
- Cloud Data Governance and Catalog Curation
- **API-enabled Data Orchestration and Provisioning**
- 3rd party data onboarding with AWS Data Exchange over 3,300 public data products

Data Producer Experience

- Package, qualify and "sell" data collections for data consumption
- Control data approval with insight into intended purpose
- **Metrics on data asset consumption, usage, and purpose**
- **Star Data Producer Gamification ratings to drive data product consumption**

Data Consumer Self-Service

- Marketplace Shopping Experience Customization
- Data Product Faceted Search and Filter for relevancy
- **Data Product Comparison Shopping**
- **API-based data consumption**

Data Producer-Consumer Collaboration

- **Real-time Interactive Chat-based Collaboration**
- Data Product Variation Hierarchy for deeper insights

CLAIRE ®

- Consumption of CLAIRE generated Intelligent Glossary Associations and Data Classifications for Data Products

Up to **90%** reduction in time-to-request for data products

400+ data products published in a few days

Up to **99%** reduction in data engineering cost for access provisioning

Data Quality and Observability

Profiling, Quality, and Exception Management

- **Data Quality natively implemented across all IDMC services**
- **OOTB data quality and stewardship experiences for full data quality lifecycle from discovery to remediation at petabyte scale**
- Open and modular DQ to support 3rd party integrations like Glue, ADF, Python Code, Hand Coding, DBT, etc.
- **Integrated, fully managed Data Exception processing for last mile data remediation**
- Discovery / Profiling across enterprise and cloud sources, open standards and common lake formats like JSON, AVRO, Parquet
- **700+ prebuilt data quality rules to accelerate DQ implementations**

Email, Phone and Address Verification

- Multi-mode address verification – batch, interactive and quick capture (auto-complete) with **Global Certifications and worldwide support**
- Global Phone Verification with standardization and Email verification with malicious threat identification

Up to **100s** of hours of manual work saved

Up to **100x** faster by leveraging AI/ML

Up to **90%** shortened implementation time with prebuilt assets

Data Quality and Observability

Data Observability

- **Automated detection of hard to find, unknown data quality issues – know the unknown**
- **Connect Data Quality metrics to business context for actionable insights**
- **Data pipeline reliability status including quality and trust metrics**
- Automated Delivery, Fulfillment & Tracking with Governed workflows in marketplace
- **Prevent data outages with proactive Detection of data quality issues and anomalies with configurable alerting**
- Continuous data observability for Outages, Critical Jobs, Connections, and Processes for predictable SLAs

CLAIRE®

- Full data quality automation to automatically assess data sources against organization standards across IDMC services
- **Profiling-based CLAIRE driven Quality Rule Generation in DQ Insights**
- **Metadata-based Data Quality Rule Recommendation and association to identify and resolve generic DQ issues**
- NLP-based Data Quality Rules generation with simple business statements to make rule building easily accessible to all

Up to **100s** of hours of manual work saved

Up to **100x** faster by leveraging AI/ML

Up to **90%** shortened implementation time with prebuilt assets

CLAIRE® AI Copilot Capabilities

Automated Data Cataloging

- Data Element Classification, **Entity Classification**, Schema Matching
- **CLAIRE-generated Classifications**

Automated Data Governance

- Intelligent Glossary Associations, Auto-DQ Rule Application
- CLAIRE Activity Dashboard

Automated Data Quality

- **Data Quality Insights** – from issue identification to remediation
- NL to Data Quality Rule Conversion

Automated Data Engineering

- Next Transformation Recommendation, Intelligent Structure Discovery, UDF/Mapplet Recommendation

Automated Master Data Management

- Identity Matching, Product Classifications, Smart File Import

Active Metadata Management

- Utilize collected metadata to provide data management automation – Join recommendations in CDI driven off Schema Matching, DQ and Mapplet Recommendations driven off data classifications and more
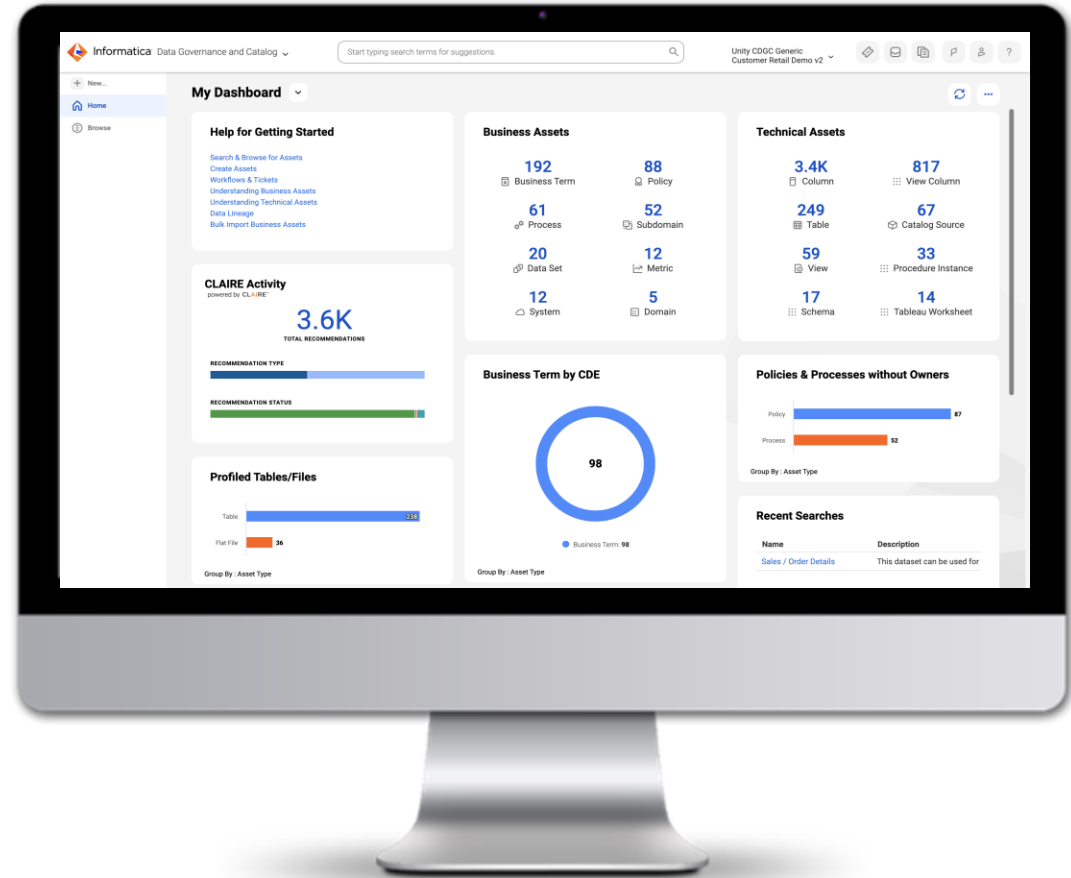- **Centralized Policy Authoring and Write-Back to select ecosystems**

Up to **95%** precision rate in automating association of business terms

Up to **$1M** in savings per month

90%+ customers use next-transformation recommendation

DEMO

Nate Rain
Data Governance Specialist
nrain@informatica.com

Donna Heimbaugh
Senior Product Specialist
dheimbaugh@informatica.com

Q&A

Where data
& AI come to LIFE

Thank You

Where data & AI come to LIFE