Informatica

29 Oct, 2024

Operationalize Risk and Compliance, Trusted Data Sharing using Data Privacy, Security and Masking

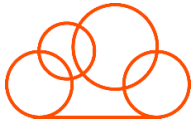
- Puneet Dudeja, Sr. Solutions Architect, CSA
- Krishnendu Chaklader, Principal Solutions Architect, CSA
- Srinivasa Gopal, Sr Principal, CSA
- Rashmi P, Sr. Solutions Architect, CSA

Where data & AI come to LIFE

Housekeeping Tips

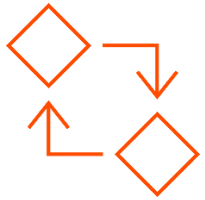
- Today's Webinar is scheduled for 1 hour
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the Q&A option and we will respond at the end of the presentation
- The webinar is being recorded and will be available on our Success Portal - where you can download the slide deck for the presentation. The link to the recording will be emailed as well.
- Please take time to complete the post-webinar survey and provide your feedback and suggestions for upcoming topics.

Feature Rich Success Portal



Bootstrap trial and POC Customers

Enriched Customer Onboarding experience



Product Learning Paths and Weekly Expert Sessions

Informatica Concierge

Tailored training and content recommendations

More Information

Success Portal

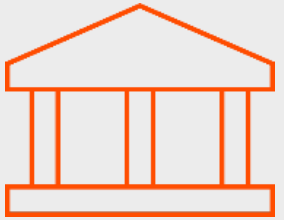
https://success.informatica.com

Communities & Support

https://network.informatica.com

Documentation

https://docs.informatica.com



University

https://www.informatica.com/in/services-and-training/informatica-university.html

Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

Agenda

1. Business Drivers and Use Cases
2. Access Management Flow - Define in one place, enforced anywhere
3. Overview of Data Classification
4. Data Democratization
5. Configure Data Access Policies
6. Policy Enforcement Policy OnRead, Policy OnLoad, Policy PushDown) - DEMO

Business Drivers and UseCases

- Data sharing and collaboration
- Cloud Migration and Modernization
- Data Provisioning
- Data for Reporting and Analysis
- Regulatory Compliance and Privacy Controls
- Mask and Protect Sensitive Data

Key trends and indicators from IDC

$952M

Estimated value of the data access market by 2027, according to the IDC analyst group

“49.3%

of data leaders indicate data governance and privacy as the highest priority initiative for organizational resources (budget, people, skills) in 2023”

Sources:
IDC, “Data Management Survey”, 2023
IDC, “Worldwide Data Integration and Intelligence Software Forecast”, May 2023

Data Privacy & Protection in CDO Insights 2024

Investment priorities

45% Data Privacy & Protection

41% Data Quality & Observability

37% Data Integration & Engineering

Data Leaders Top Responsibilities

30% Data analytics & insights

29% Data privacy, protection and compliance

29% Data strategy & governance

28% Improving data literacy & data culture

27% Enabling stakeholder collaboration

26% Enabling data sharing & democratization

Challenges adopting GenAI

42% Quality of Data

40% Data Privacy & Protection

38% AI Ethics

Gartner Perspective: Data Security Platforms

Macro-trend

"By 2025 30% of Gartner clients will protect their data using a "A need to share" approach rather than the traditional "Need to know" approach."

Organizational need

"Organizations use an increasingly complex set of security controls. Successful SRM leaders can significantly improve business utilization and data value by building a migration plan from siloed data security offerings to data security platforms enabling simpler, consistent end-to-end data security.

Category definition

"Data security platforms (DSPs) aggregate data protection requirements across data types, storage silos and ecosystems, starting with data discovery and classification. DSPs typically protect data by using late binding access controls, for example data masking, format-preserving encryption (FPE) or tokenization."

Sources:
Gartner, "2022 Strategic Roadmap for Data Security Platform Convergence", September 2021
Gartner, "Hype Cycle for Data Security, 2022", August 2022

Modern Data Governance

Ensure data availability, reliability and governance

Manage risk and comply with regulatory mandates and policies

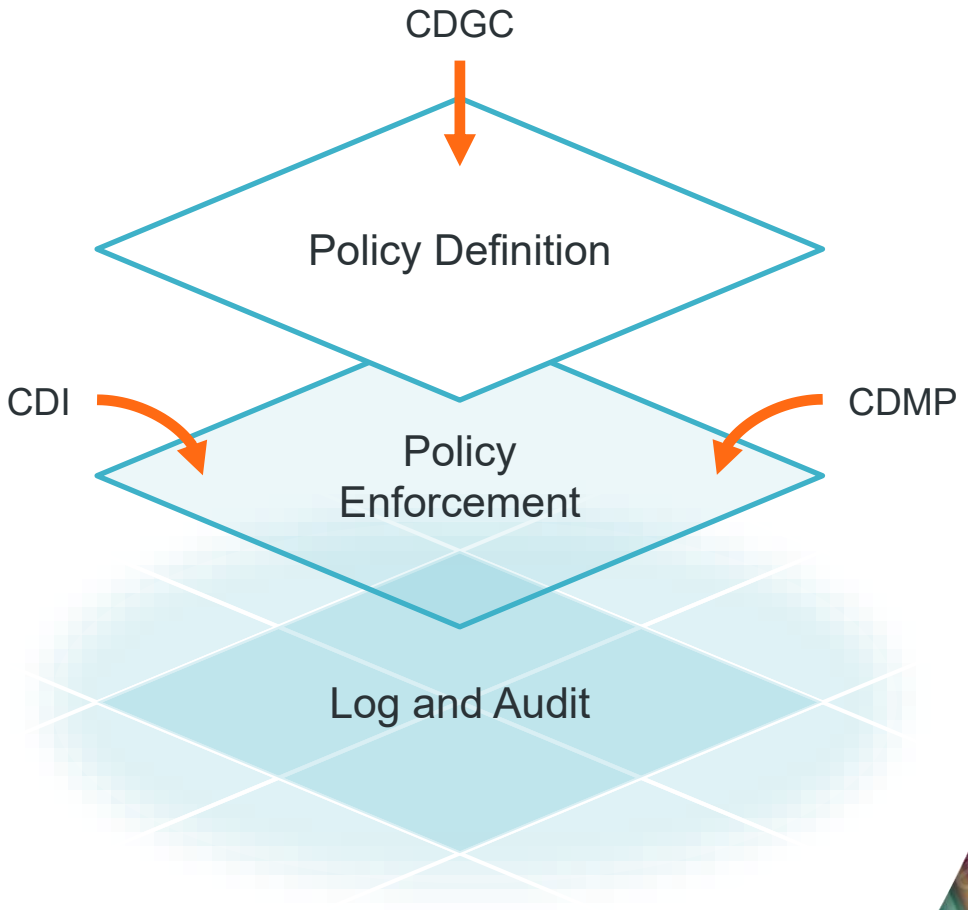
Open data for business with actionable insights and automation

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

Informatica's Cloud Data Access Management

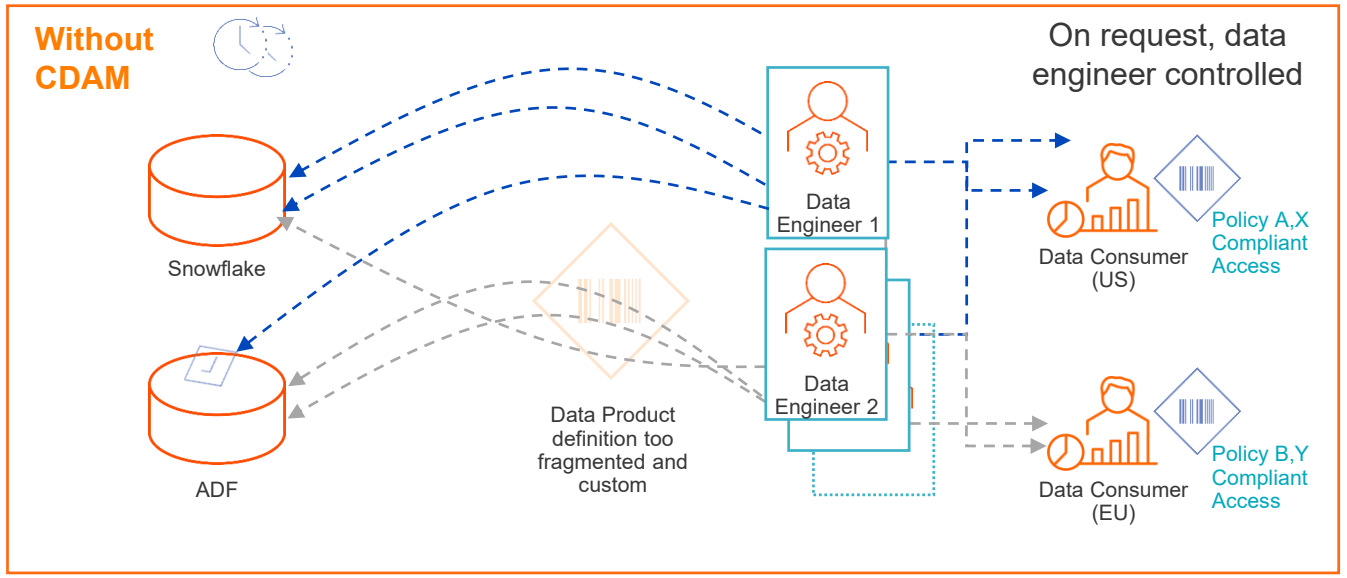
Ensure the right people have access to the right data at the right time

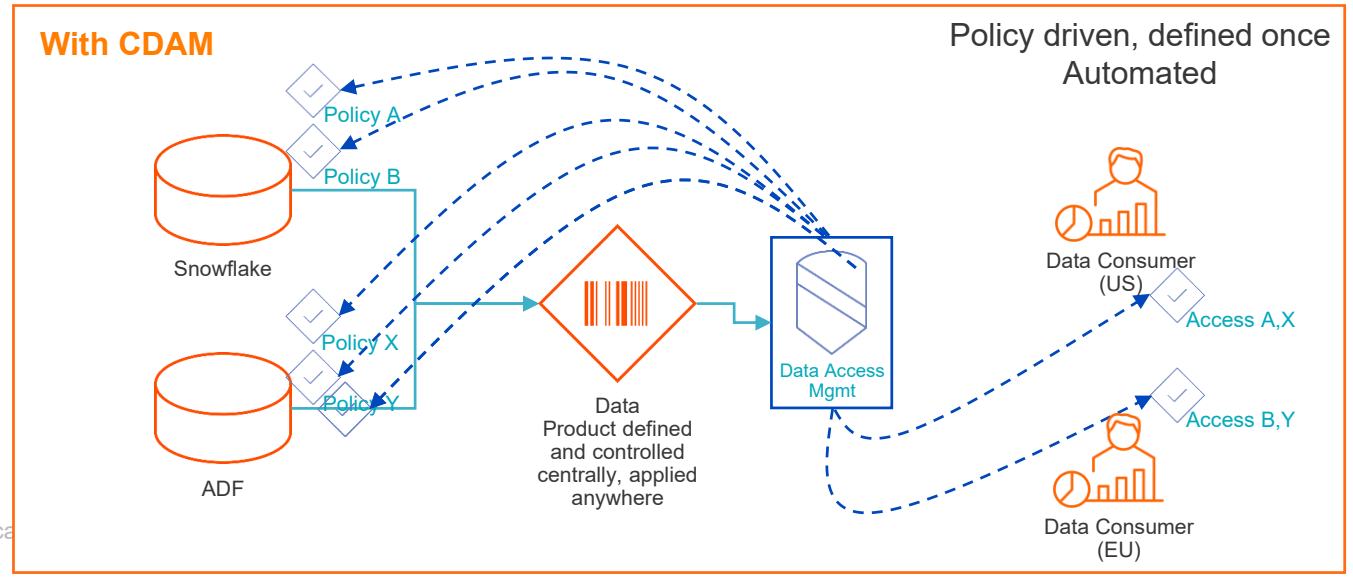
- Logical, scalable, metadata-driven policies
- Leverage discovery and classification from CDGC
- Defined in one place, enforced everywhere
- Broad range of policy enforcement points, including within CDI workloads
- Hybrid- and multi-cloud enforcement
- Streamline access requests in CDMP
- Observe, monitor, and audit data access



What is Data Access Management?

- **Data Access Management** enables organizations to reflect their data strategy with fine-grained policies and enforce controls for appropriate access to business-critical, sensitive data.
- These controls consider the conditions of the operating environment, the user and the application to ensure data use is appropriate and limits risk exposure by aligning with compliance rules.
- Automated data security and privacy policy enforcement, underpinned by data discovery and classification, metadata integration, and scalable conditional logic for contextual accuracy, accelerates safe and trusted access and use of data that is fit for purpose, accurate, and reliable to build consumer confidence in achieving expected business outcomes.

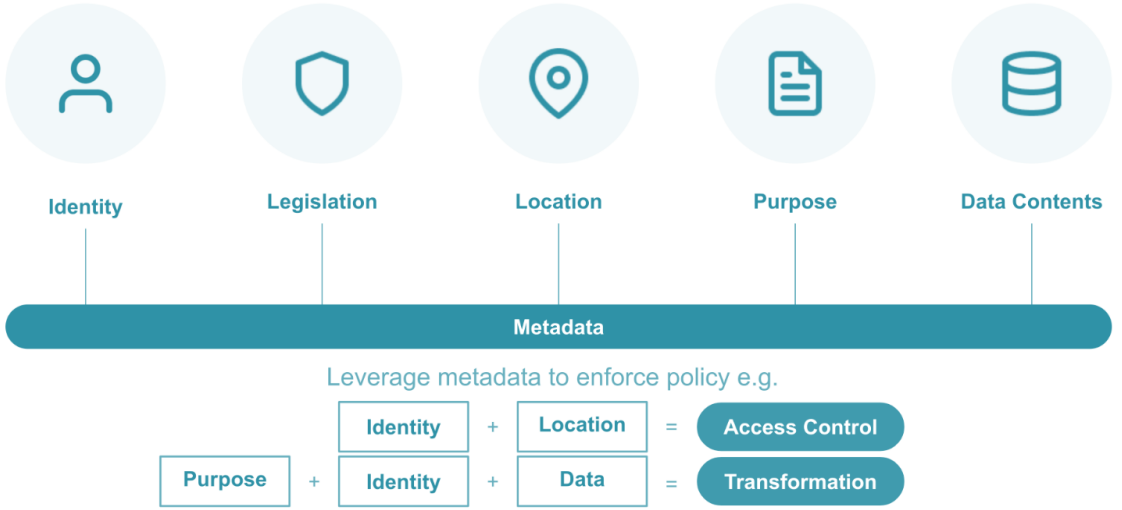




Informatica

CDAM - Automated policy driven data access and controls

- Consistent enforcement across all data products, based on context and user
- Multi cloud, hybrid (on premise and cloud), single cloud = deploy once, enforce everywhere
- Robust privacy controls
  - ABAC
  - Apply masking techniques (tokenization, encryption, hashing and more advanced generalization, substitution, k-anon type controls)
  - Provide re-identification based on certain context information – user identity, role, purpose, etc.
- Auditing and logging to evidence policy enforcement

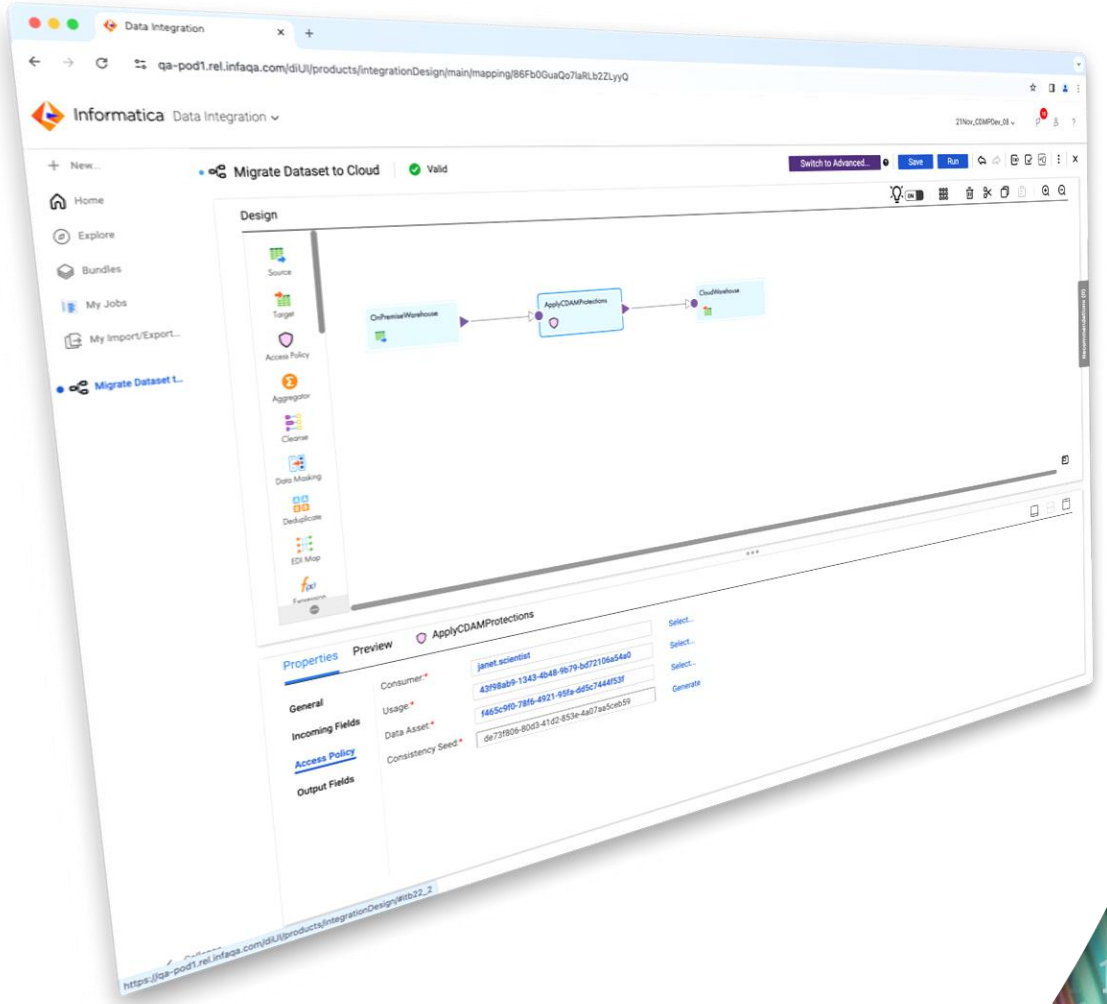


CDAM for Data Migration

Implement rigorous security and privacy postures across cloud platforms, on premises and hybrid

Cloud Data Integration (CDI) moves data between stores, including from on premises to the cloud.

Use CDAM to land de-identified data in the cloud, protecting sensitive data in structured data sets.



Supports key initiatives:

- Digital transformation and modernization
- Migration to the cloud
- Data for AI and ML training
- Regulatory compliance

CDAM for Data Democratization

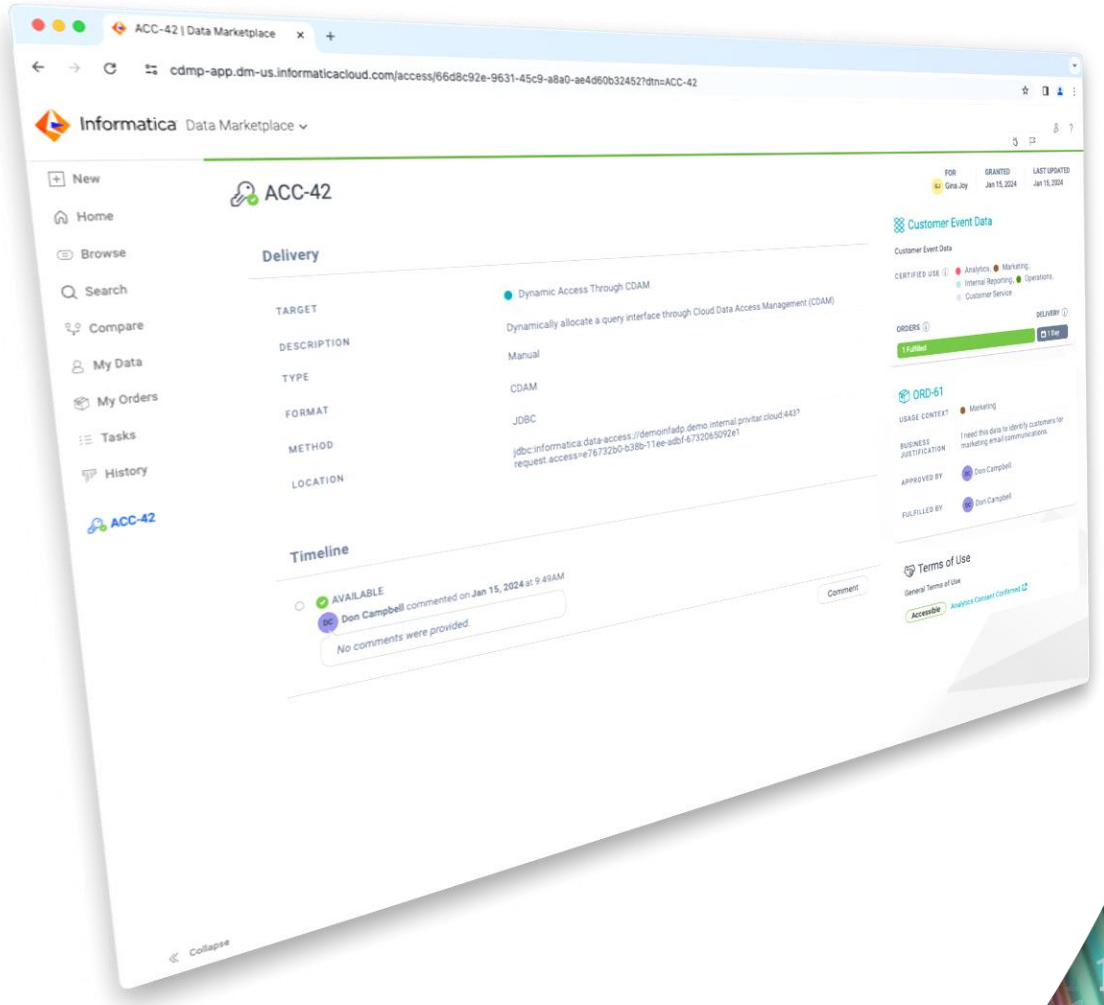
Extend data access to more and more users and applications through self-service

Cloud Data Marketplace (CDMP) provides self-service and data shopping user experiences.

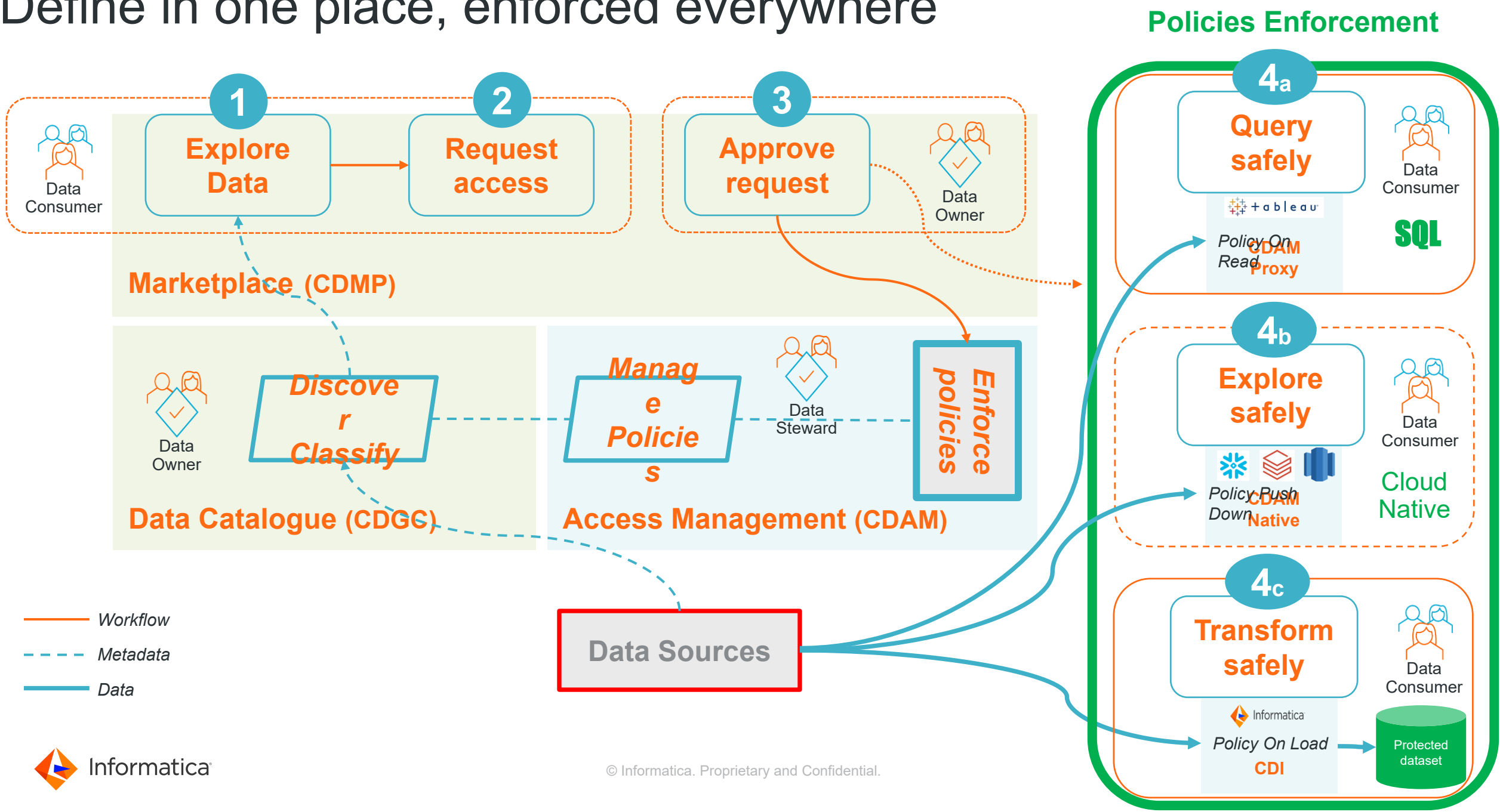
CDAM dynamically de-identifies sensitive data in the cloud using in-store controls or a proxy.

Support key initiatives:

- Data marketplace and shop for data
- Data sharing and collaboration
- Data for AI and ML training
- Regulatory compliance



Define in one place, enforced everywhere



Informatica

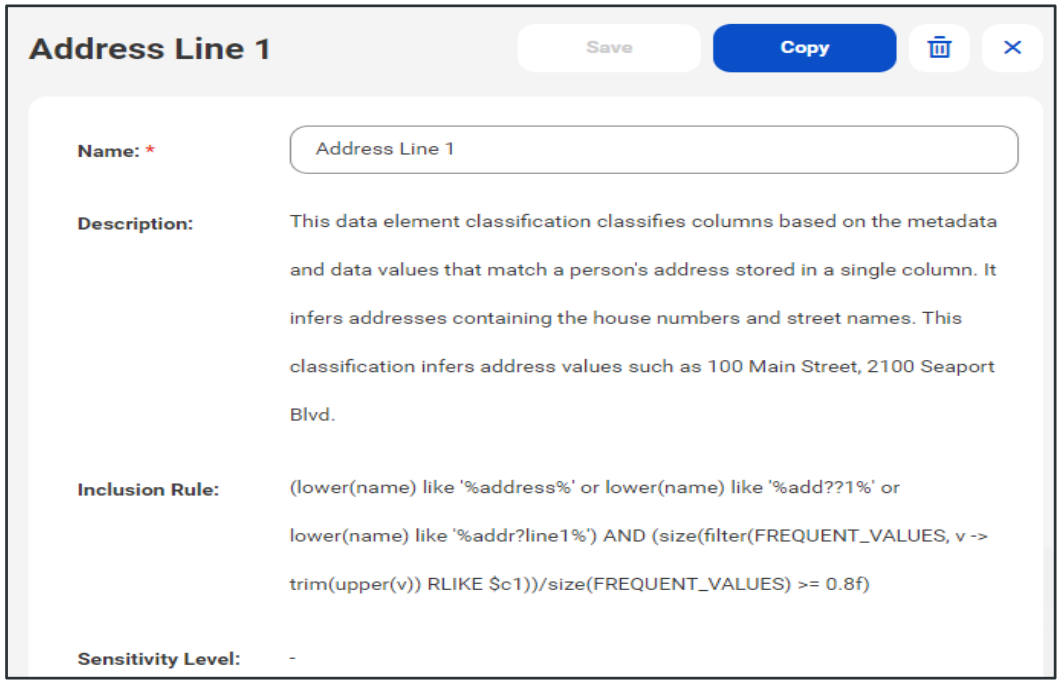
Data Classification

Data Classifications

Means to identify and organize data into categories based on their functional meaning to add more context in the catalog and tag sensitive data as well.

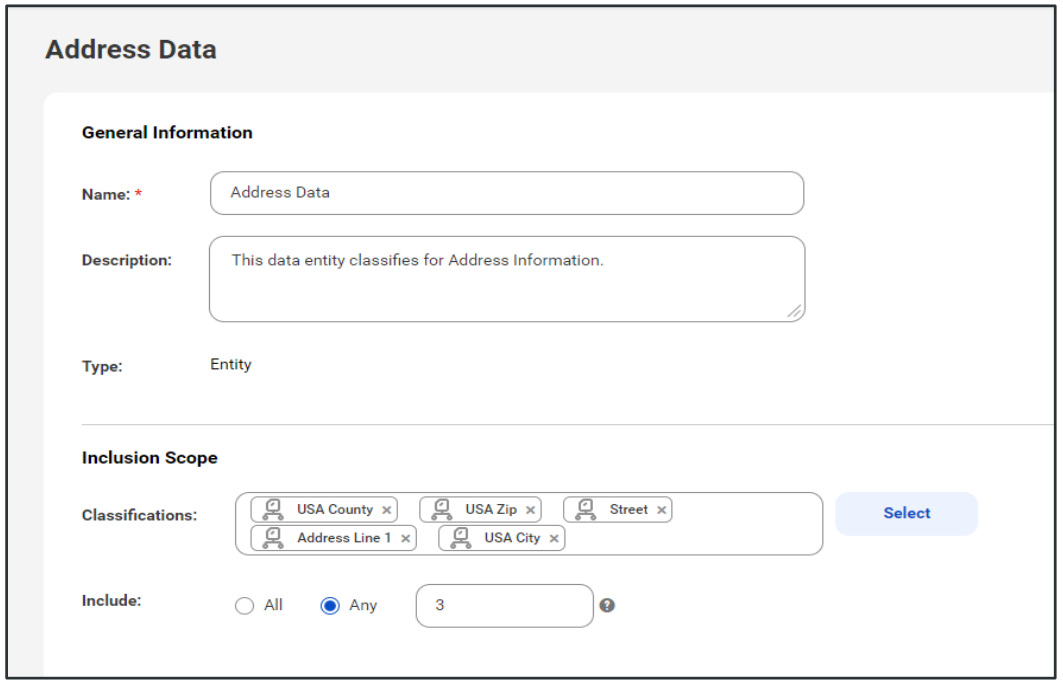
Data Element Classifications

To classify an individual attribute within a table or file based on the type of data stored. These are generally rule based classifications which can run on metadata and data both. Example Email, SSN, Address Line 1 etc.



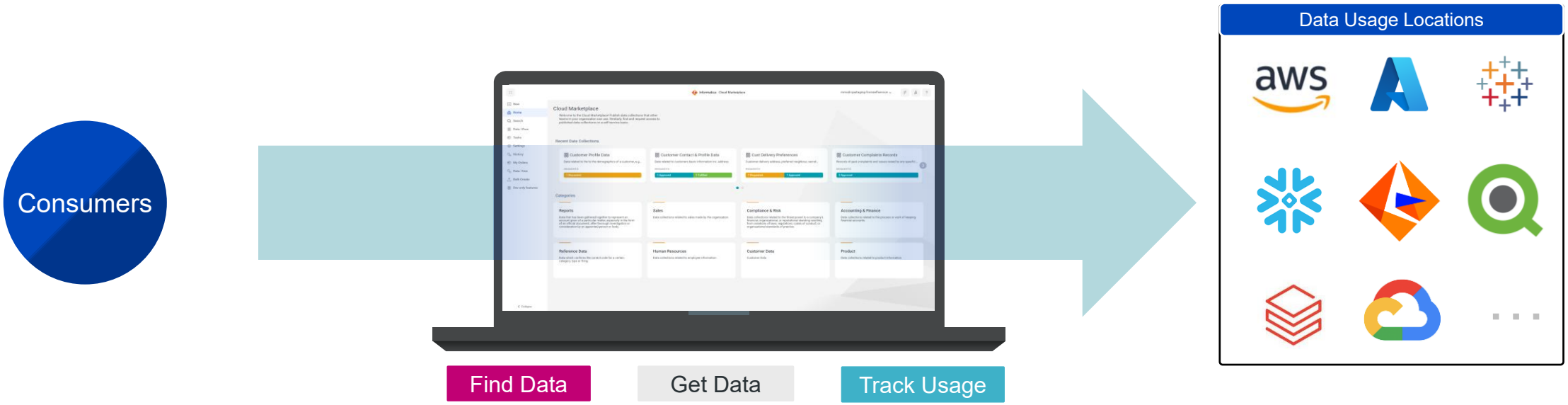
Data Entity Classifications

To classify an entity like purchase order, person, Address with the help of a group of attributes present in the record within a table/file. For Example, Emp ID, Name Email, DoB, Phone can help classify a table to contain Employee information.

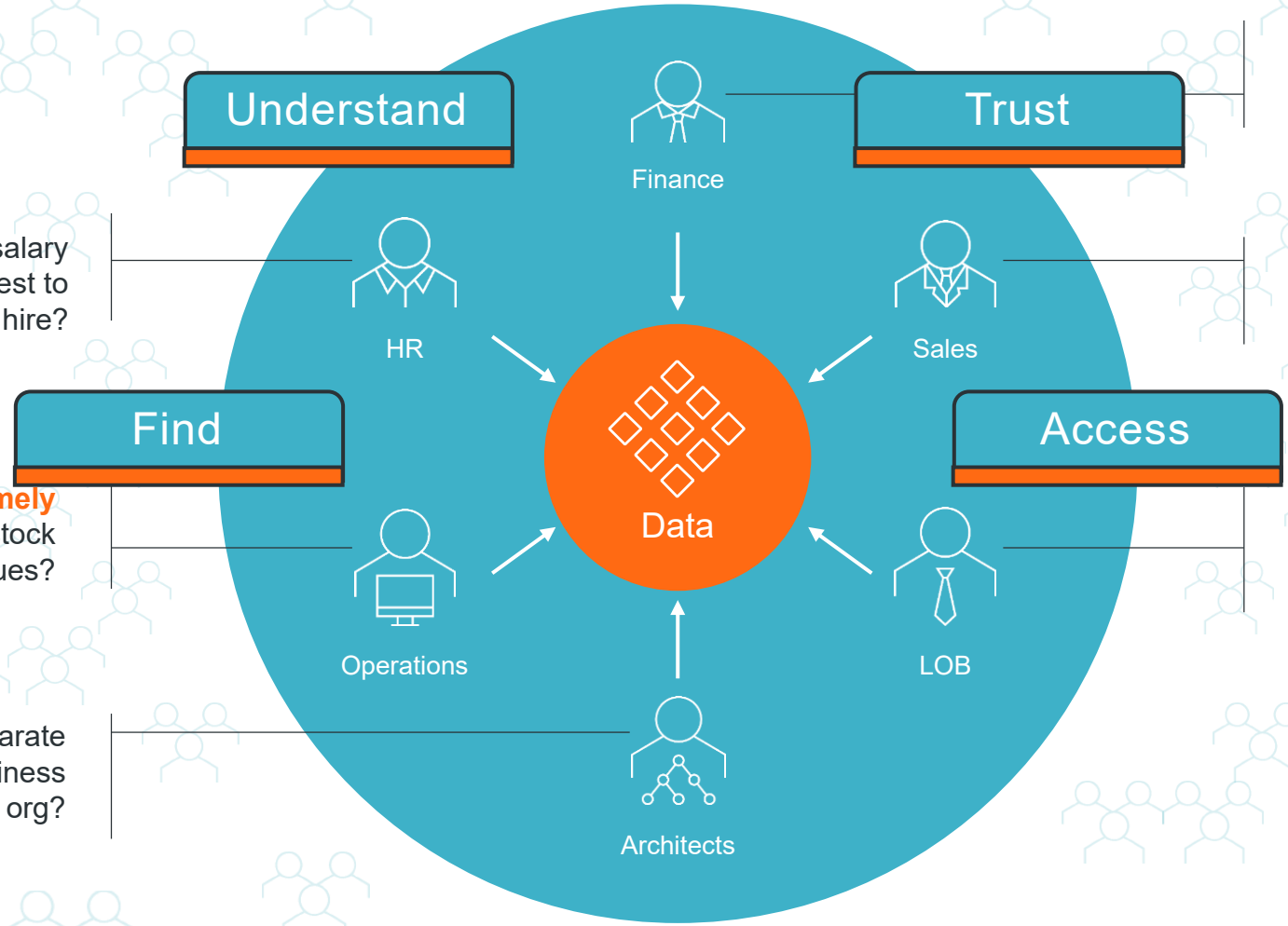


Data Democratization using Cloud Data Marketplace

Every consumer, every source, every destination



Lifecycle of Data Democratization



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

Configure Data Access Policies (CDAM)

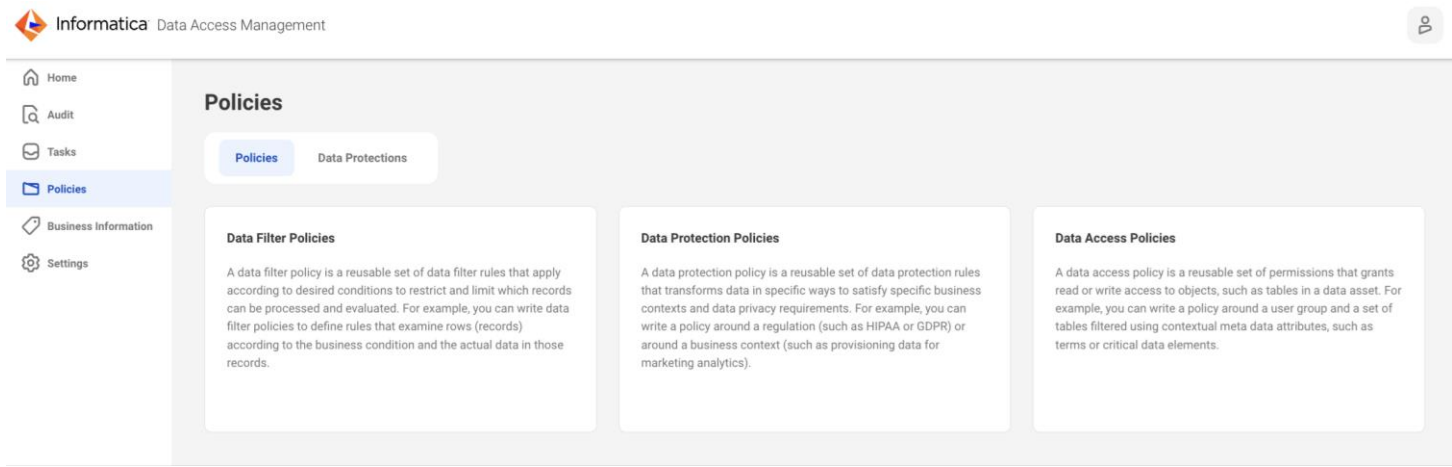
Data Access Management

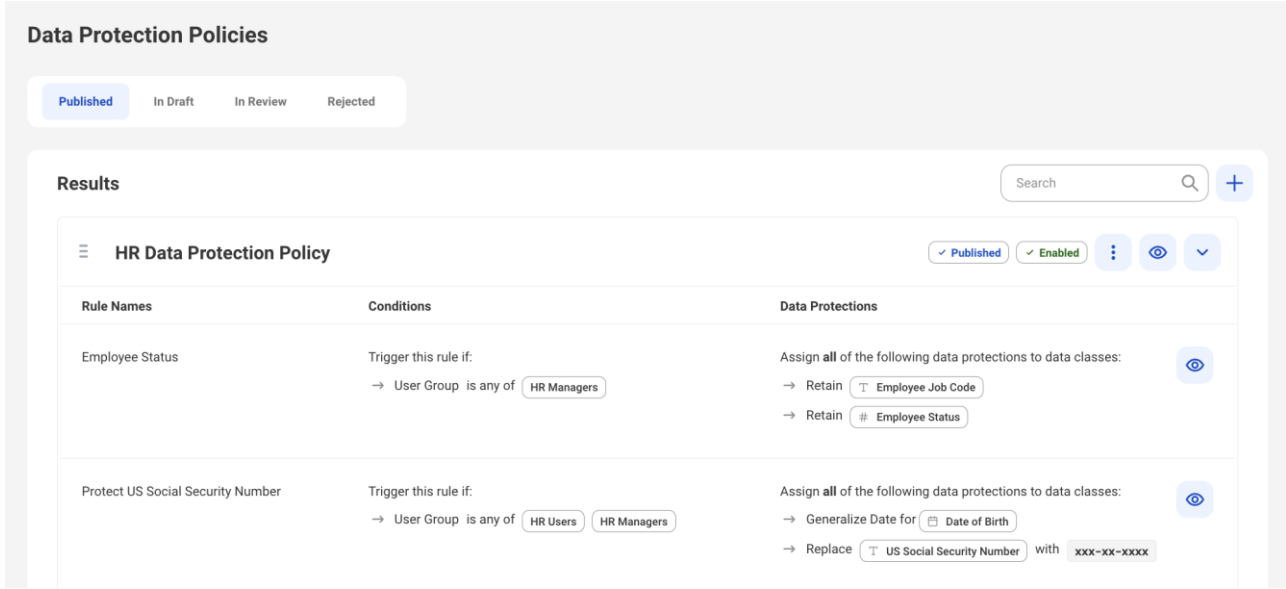
Policies represent a set of Rules that apply protection for common set of requirements or standards.

Policies can allow or deny access to data, based upon the content of the data.

Rules contain business logic that define the conditions of when to apply actions to Data Assets and their elements.

Rule actions apply different transformations to data elements to protect them, and can apply different transformations based upon the content of a single data element





Configure Data Access Policies (CDAM)

Sensitive value	*Transformation (PET)*	*Protected value*	*Description.*
C. Mason	**Retain**	C. Mason	The input value is unchanged. The output value matches the input value.
C. Mason	**Redact with NULL**	*NULL*	The output value is replaced with NULL
C. Mason	**Constant Text**	*****	Replace all input values with the same user-defined value.
C. Mason	**Truncate**	C.	Remove or retain part of the input value.
504-CD-968Y	**Tokenize Text**	363-GT-047A	Replace input values with randomly generated text matching a user-defined regular expression.
1,456.78	**Tokenize Number**	98,263.21	Replace input values with a randomly generated number that matches a user-defined regular expression.
1991-06-24	**Generalize Date**	1991-01-01	Replace the input value with a generalized version of the date or a constant date if the value is outside of a user-defined date range.

Option to provide consistent protected values
If a sensitive value appear multiple times, it will be masked with the same protected value.

Policy **Enforcement**

Policy management on top of data classification

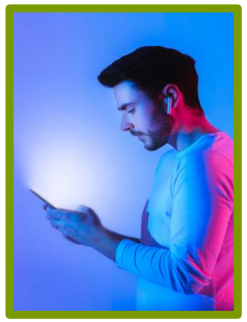
Policy enforcements OnLoad OnRead PushDown

OnLoad

Demo2

Cloud Migrations & Modernizations

Conrad



OnRead

Demo3

Data Democratization

Bob

Gina

PushDown

Demo4

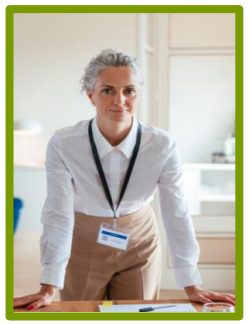
Native controls

Clara

Martin

Demo1

Central Data Policy Governance

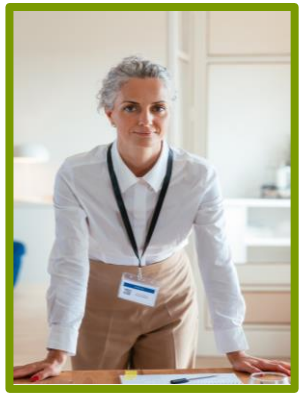


Grace

Demonstration

Story #1: Central data policy governance with Grace

Data Steward



Challenges

Ensure all data access requests are managed by data policies.

Today:

- She needs to work with too many Data Owners managing distinct data storage solutions
- How to define central policies that could be used in all data usage topics?

Policy **management** on top of data classification

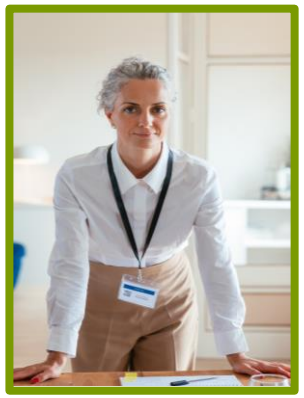
Demo1

Central Data Policy Governance

Central data policy governance

Cloud Data Access Management from Informatica IDMC

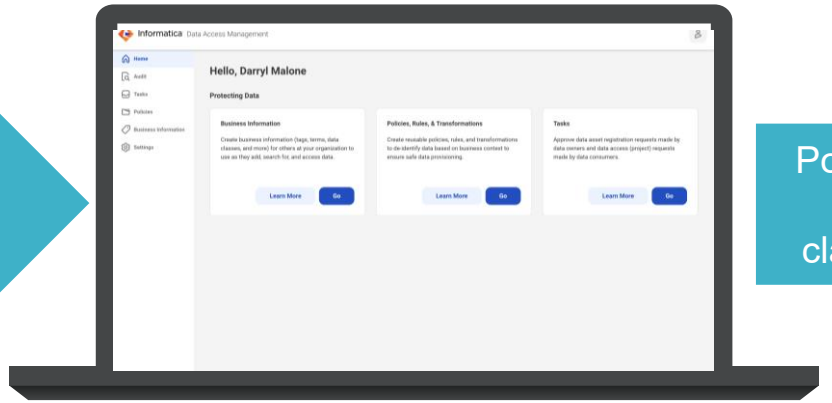
Grace



Data Steward

Define data policies

CDAM



Policies on top of data classifications

- **Filters**
- **Protections**
- **Audit**

Classifications

CDGC

Sensitive data are protected and comply with global privacy regulations including CCPA and GDPR.

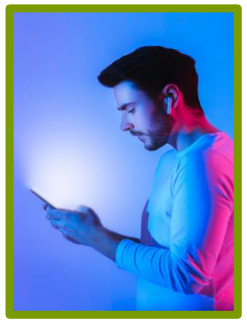
No need to replicate policies in several system.

Solutions

Demonstrations

Cloud Migrations & Modernizations

Conrad



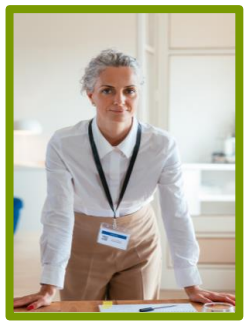
Demo2

Cloud Migrations & Modernizations

OnLoad

Policy management on top of data classification

Policy enforcement s OnLoad OnRead PushDown



Grace

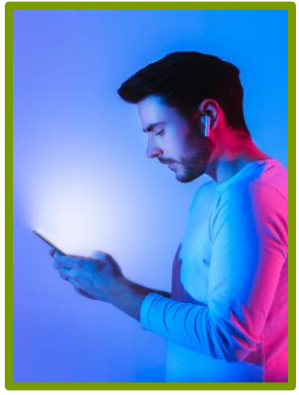
Demo1

Central Data Policy Governance

Demonstration

Story #2: Cloud migration & modernization with Conrad

Data Scientist

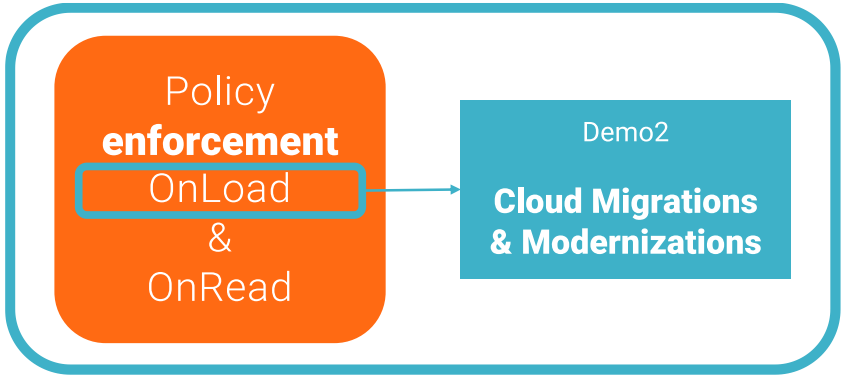


Challenges

Conrad needs employee data in the Cloud to use his AI tools

Today:

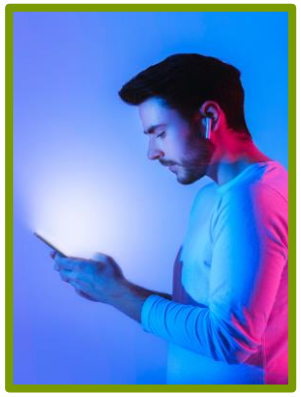
- How could Conrad get safe datasets in the cloud?
- How to reduce the pain of the Data Engineer who should design a data mapping?
- How to keep data utility and consistencies from distinct data sources?



Cloud Migrations & Modernizations

Cloud Data Integration from Informatica IDMC

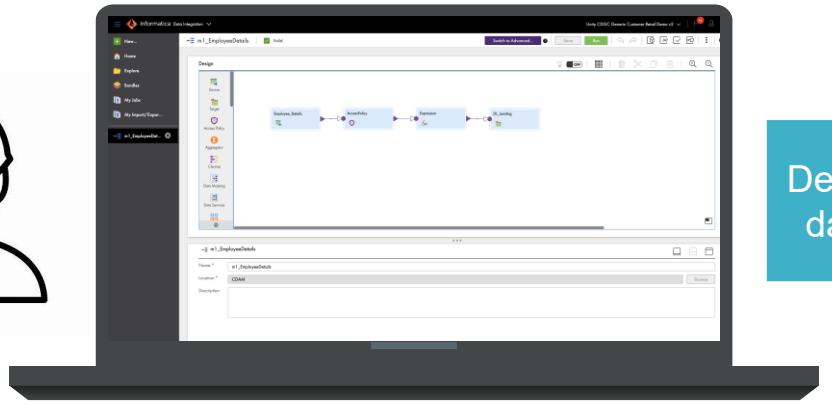
Conrad



Data Scientist

Request to Data Engineer

CDI



Deliver protected data anywhere

Protected dataset

Data Sources

Data Lake

Conrad gets datasets on Cloud even from distinct sensitive sources

Data engineer don't have to know how to protect which data.

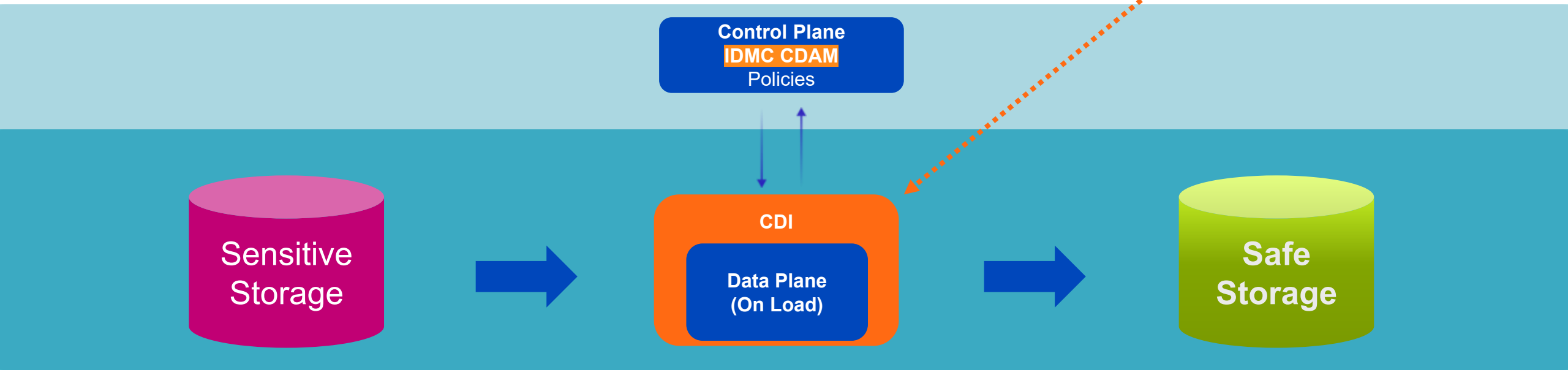
Conrad will get consistent protected dataset keeping data utilities

Solutions

Policy enforcement: On Load

Static protection

- Get a protected physical asset from a sensitive one
- Execute in CDI using the stage “Access Policy”



Demonstrations

Data Democratization

Policy management on top of data classification

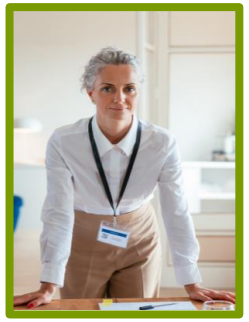
Policy enforcements OnLoad
OnRead
PushDown

OnLoad

OnRead

Demo1

Central Data Policy Governance

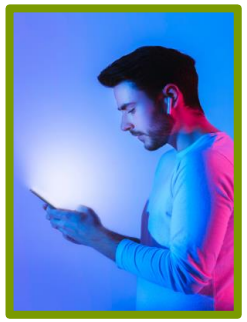


Grace

Demo2

Cloud Migrations & Modernizations

Conrad



Demo3

Data Democratization

Bob

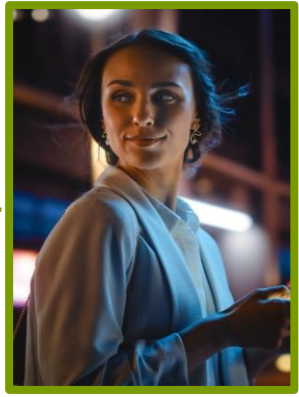
Gina

Demonstration

Story #3: Data democratization

HR Manager (USA)

Business Analyst



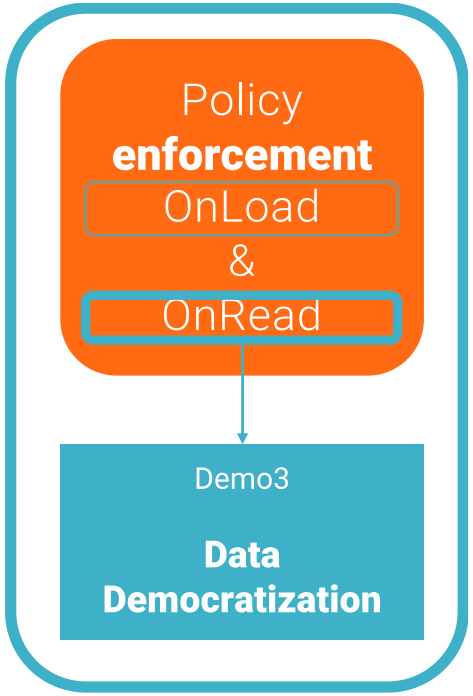
Challenges

Bob needs a report on his employees. He's based in USA.

Gina needs to analyze employee salary all around the globe using her favorite SQL query tool, Tableau.

Today:

- Gina & Bob spend too much time explaining their projects to too many approvers.
- How to reduce time to data keeping data utility?



Data democratization

Cloud Data Marketplace from Informatica IDMC

Bob

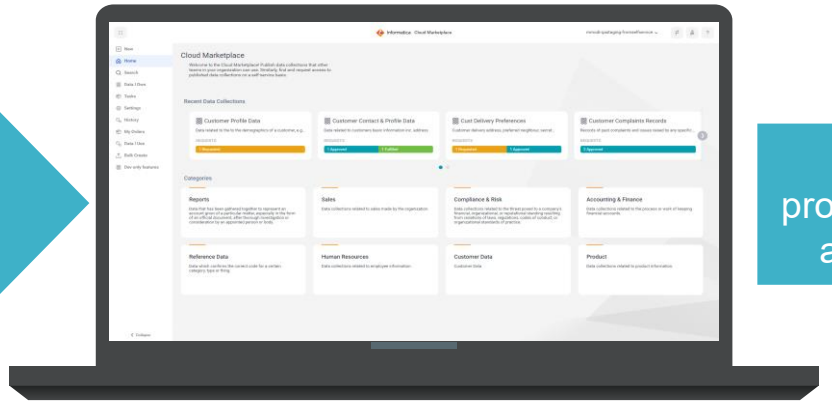
Manager

Gina

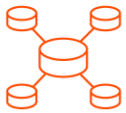
Analyst

Shop for data

Informatica Cloud Data Marketplace

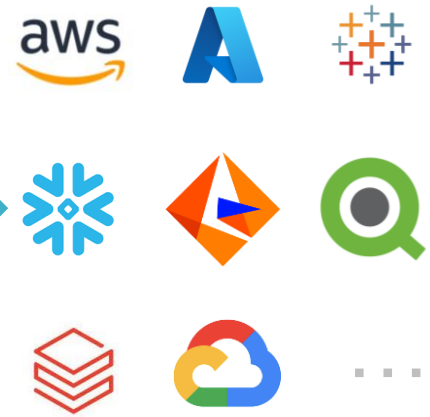


Data Products



Deliver protected data anywhere

Data Usage Locations



Bob needs details about his employees.

Gina require employee and salary data for analytics.

Gina & Bob know where to order access to data.

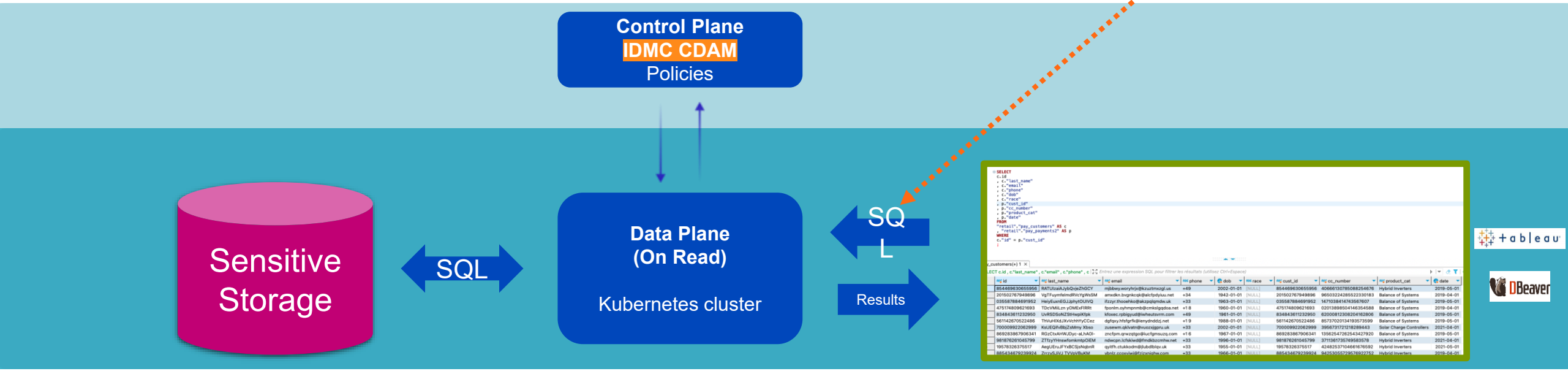
They can run their analysis with their favorite tool.

Solutions

Policy enforcement: On Read

Dynamic protection

- Automatic protection on SQL query results
- Data access management **Proxy**



Demonstrations

Native controls

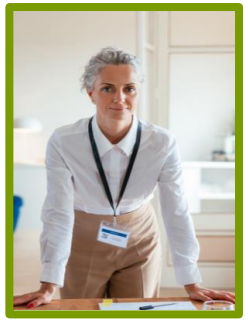
Policy management on top of data classification

Policy enforcement s OnLoad OnRead PushDown

Demo1

Central Data Policy Governance

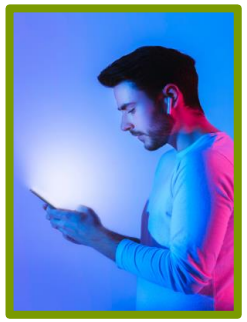
Grace



Demo2

Cloud Migrations & Modernizations

Conrad



Demo3

Data Democratization

Bob

Gina

Demo4

Native controls

Martin

Clara

Informatica

Demonstration

Story #4: Native controls

HR Analyst

HR Manager

Challenges

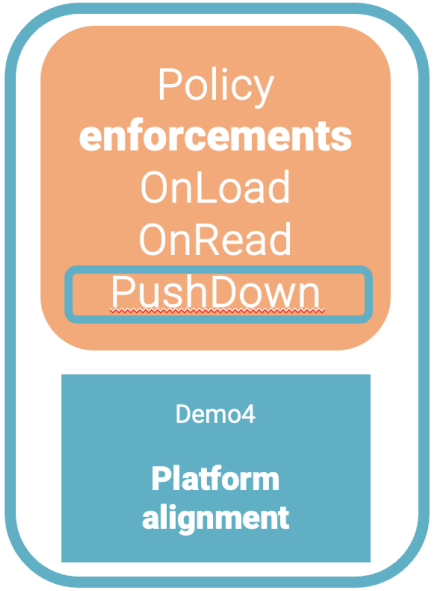
Snowflake user access data only in Snowflake User Interface or Dashboard.

Other access it is using Python.

How to apply CDAM policies into Snowflake?

Today:

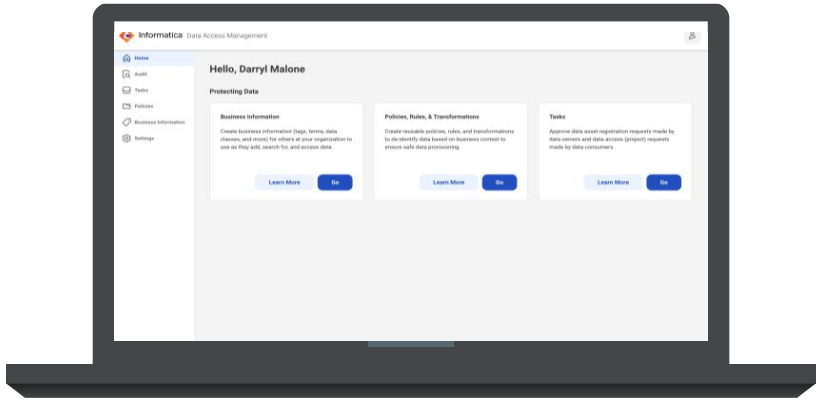
- IT needs to configure access policies into Snowflake platform additionally to the other platforms



Native controls - Snowflake

Cloud Data Access Management from Informatica IDMC

Informatica Cloud Data Access Management



Synchronize data policies

* Native Controls on Databricks on 2024.11 roadmap

Some user access only to snowflake

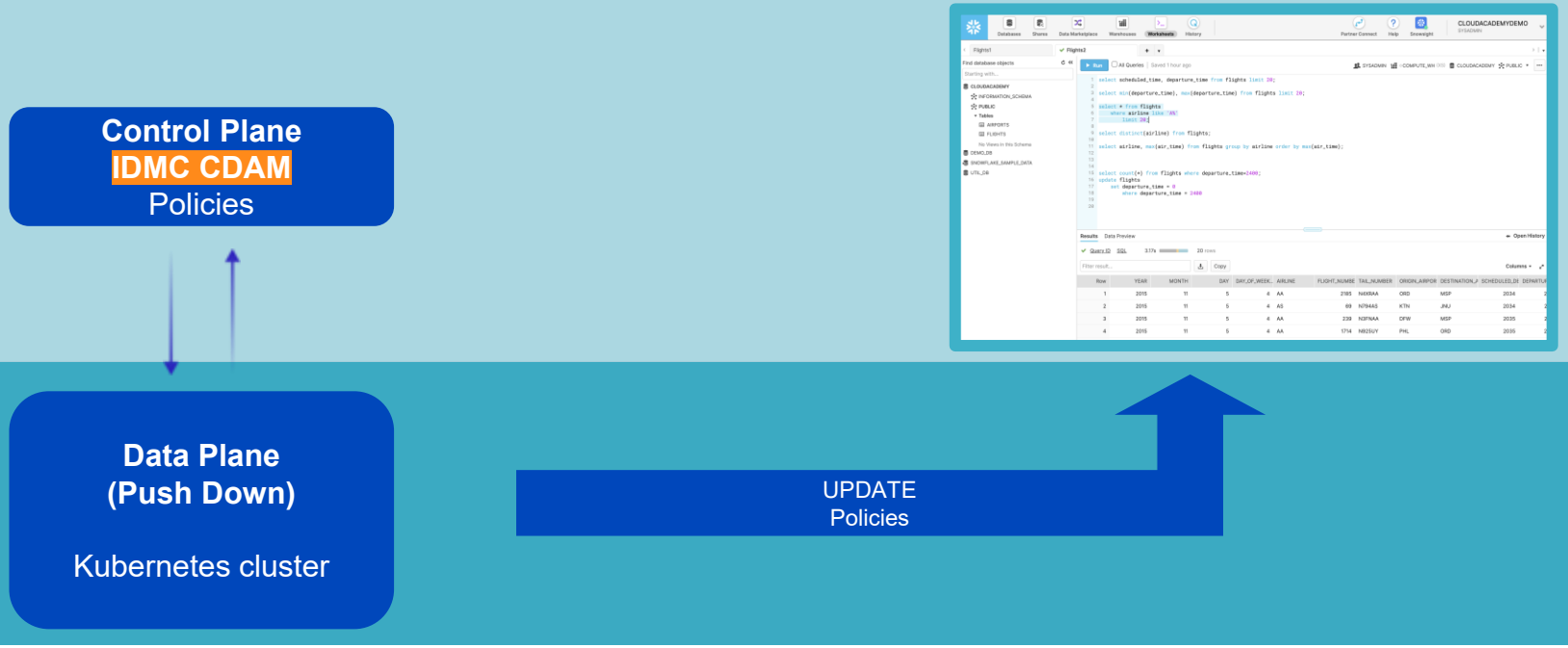
Snowflake users benefit from CDAM policies

Solutions

Policy enforcement: Push Down - Snowflake

Native controls

- Synchronization engine manage in Kubernetes
- Update INFA policies into Snowflake



References :

- Informatica Cloud Data Marketplace -Introduction and Getting Started
- TT Webinar - Data Discovery Best Practices in Cloud Data Governance and Catalog
- TT Webinar - Policy Driven Controls in protecting Sensitive & Confidential Data
- TT Webinar - CDMP Order Fulfilment and Automated Provisioning Use Case Study
- TT Webinar - Enabling Self-Service Analytics using CDMP, CDGC and CDQ
- TT Webinar - Data Observability in Cloud Data Governance and Catalog
- TT Webinar - Rest Accelerator Pack - Cloud Data Marketplace
- Experience Lounge - IDMC

Informatica

Thank You

Where data & AI come to LIFE