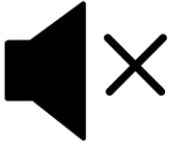Informatica®

23 July, 2024

Unleashing the Power of Search in CDGC

- Giri Chintarlapalli, Solution Architect, IPS

- David Gaffaney, Sr Principal, IPS

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.
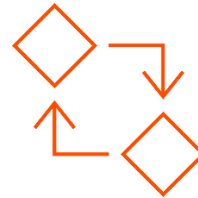
Feature Rich Success Portal
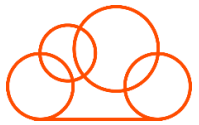
Bootstrap trial and POC Customers

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
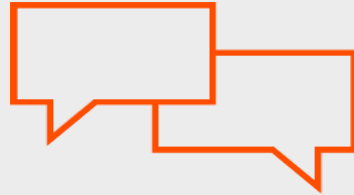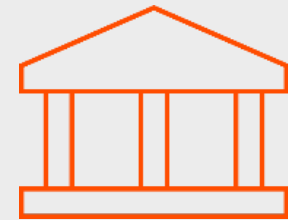
Informatica®

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.

Our Speakers

Informatica Professional Services

David Gaffaney

Advisory Services Sr Principal

Giri Chintarlapalli

Solution Architect

Today's Agenda

1 Introduction and Housekeeping

2 Cloud Data Governance and Catalog

3 Facets and Other Key Concepts

4 Basic Navigation for Consumers

5 Tiers of Advanced Search Techniques

6 Demonstration

7 Q&A

8 Close

Setting the Stage

This webinar is intended for **Data Stewards** and **Business users** in a **Data Consumer** Role

We will explore advanced methods of Finding **Data Catalog and Governance** information in **CDGC**, using standard and advanced searches.

We hope to teach you **innovative search techniques** to help you navigate and target your essential Catalog and Governance Data in support of **business use cases**.

Disclaimer
This is not a ClaireGPT Webinar, Stay Tuned for the next deep-dive.

… so you can focus on your **Core Business** and support **Self-service Data Discovery** for technical and non-technical users.

Informatica

What are your Goals?

- If we think about common **Consumer Use Cases** we can get a sense of what our users want to get out of the Informatica Model.

- This is all about enabling both our **Technical** and **Business Stewards**.

Data Self Service

"I'm looking to develop a dashboard for the Underwriters and I need sources of Claim Payment Data."
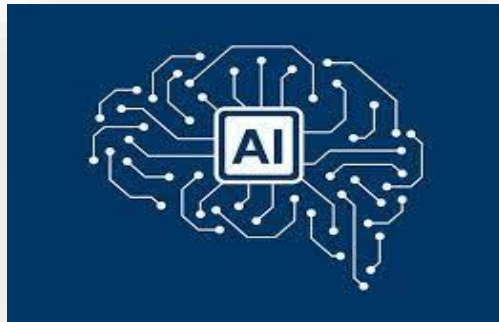
Feeding AI Models

"I am looking for the right data sets to feed the Customer Service LLM."

Modernization

"We are looking to move key on-prem data to a SaaS solution and need to investigate the integration points."

Governance and Metrics

"I need to validate that Critical Data Elements that feed the KPIs are of high quality and traceable."

IDMC: Cloud Data Governance and Catalog in Context

Facets Model in CDGC

- This is a high level ERD showing the Facets within the CDGC Model

- Today's Focus is primarily on the Catalog Data component, with references to Business Impact.

Business Impact and Change captures policy and process context along with managing evolution of the model

Project
Rep... bo...

Business Area
Depar... wi... res...

Geography
Repres... broa... ty... regu...

Legal Entity
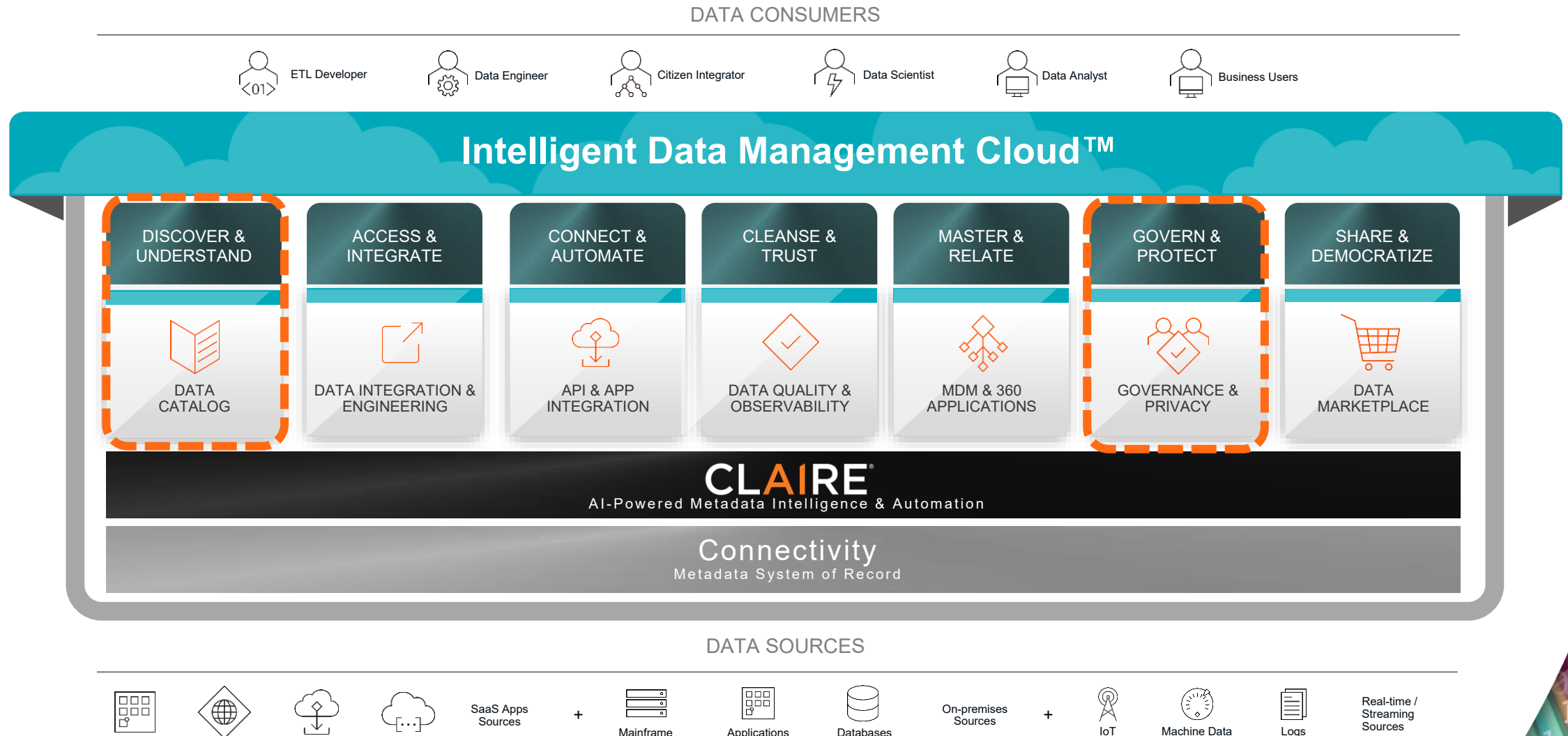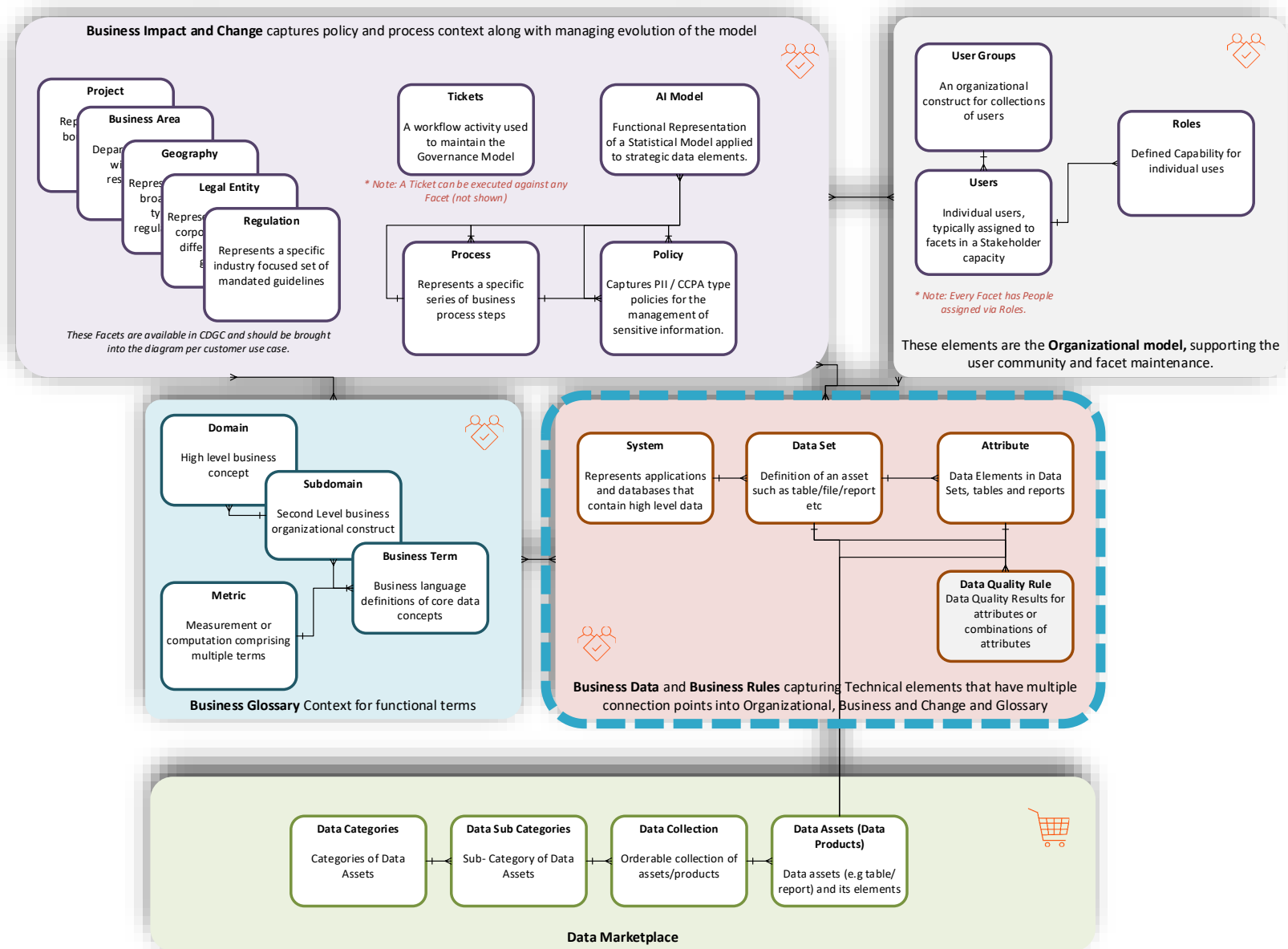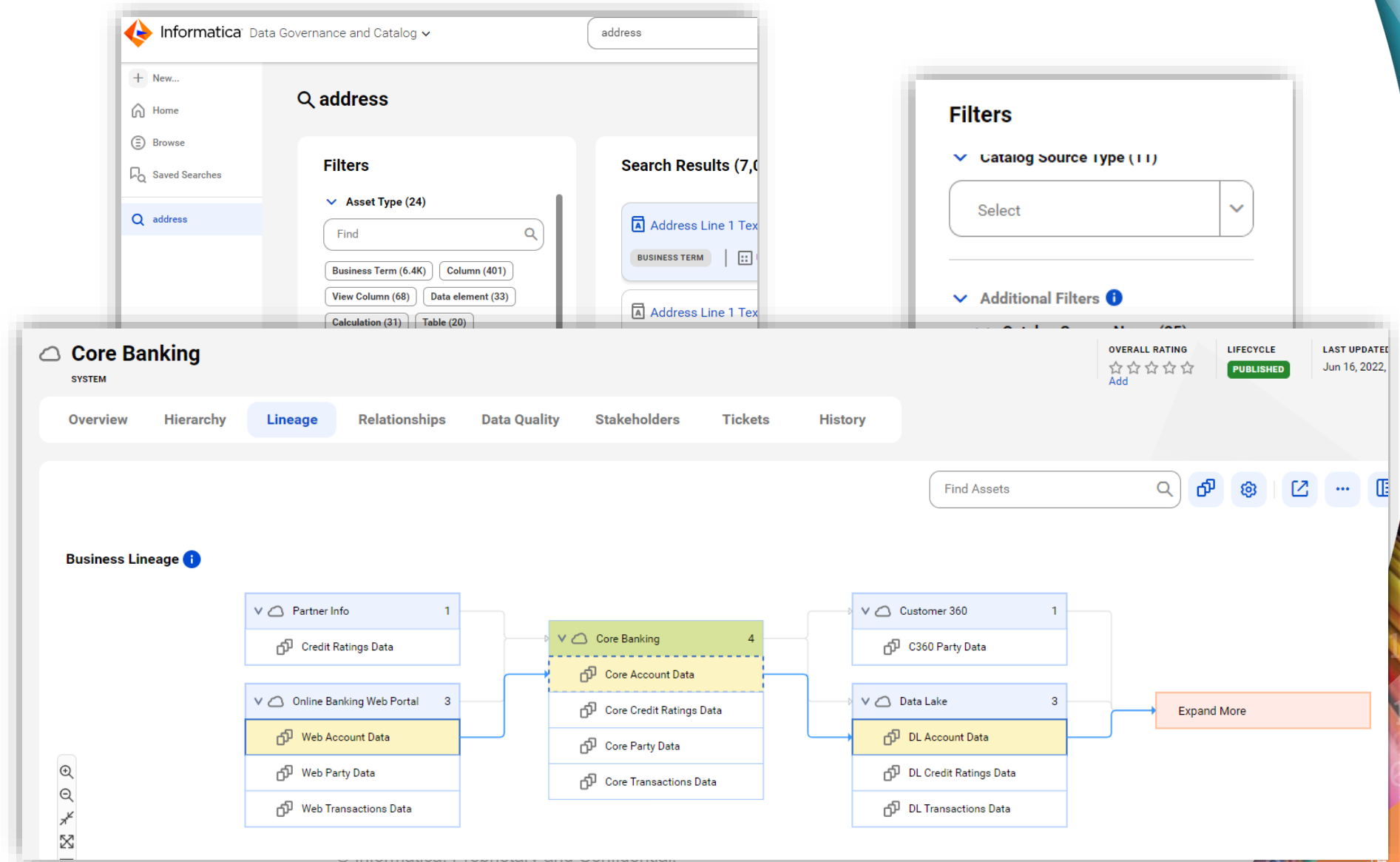Repres... corpo... diffe... 8...

Regulation
Represents a specific industry focused set of mandated guidelines

Tickets
A workflow activity used to maintain the Governance Model

AI Model
Functional Representation of a Statistical Model applied to strategic data elements.

Note: A Ticket can be executed against any Facet (not shown)

Process
Represents a specific series of business process steps

Policy
Captures PII / CCPA type policies for the management of sensitive information.

These Facets are available in CDGC and should be brought into the diagram per customer use case.

User Groups
An organizational construct for collections of users

Roles
Defined Capability for individual uses

Users
Individual users, typically assigned to facets in a Stakeholder capacity

Note: Every Facet has People assigned via Roles.

These elements are the **Organizational model,** supporting the user community and facet maintenance.

Domain
High level business concept

Subdomain
Second Level business organizational construct

Business Term
Business language definitions of core data concepts

Metric
Measurement or computation comprising multiple terms

Business Glossary Context for functional terms

System
Represents applications and databases that contain high level data

Data Set
Definition of an asset such as table/file/report etc

Attribute
Data Elements in Data Sets, tables and reports

Data Quality Rule
Data Quality Results for attributes or combinations of attributes

Business Data and **Business Rules** capturing Technical elements that have multiple connection points into Organizational, Business and Change and Glossary

Data Categories
Categories of Data Assets

Data Sub Categories
Sub- Category of Data Assets

Data Collection
Orderable collection of assets/products

Data Assets (Data Products)
Data assets (e.g table/ report) and its elements

Data Marketplace

Basic Navigation Strategies

Think of these as Tier Zero browsing

- Top Down, Filter Selection

- Navigating Impact Maps

Advanced Search Techniques

Tiers of Search Insight

Tier 1

- **Basic Searches** using Exact Text
- Used for very specific searches when you know your **exact target**

Tier 2

- **General Searches** based on source type or other broad category
- Supported by **Filters** as needed.

Tier 3

- **Complex Searches** using specific syntax that supports narrowing down hits when there are very **Large Result Sets**

CDGC – Advanced Searches

Search queries

Phrase or Combinations of keywords

Find asset types, asset names

Intelligent autofill

Informatica

Privileges

Standard roles

Data Governance User

Custom roles

View Technical assets
Read permission for "Technical Assets" type

Informatica

Common search terms

Exact searches

- Use single quotes or double quotes
  - Examples: 'Net Profit', "GDPR-Violations"

Partial terms

- Wild card entry *
  - Examples: catalog source Amazon*, 'phone num*

Multiple search queries

- Use comma (,)
- OR conditions
- Examples: domain 'Party Data', subdomain Measures

Connected assets

- Use parentheses ( )
- Search for connected assets
- Search query contains a predicate
- Example: (column with name Customer) related to table

Common search terms - Continued

Synonyms

- Alternative names in the search box
- Example: Policies, Policy; Columns, Data Elements related to table

With keyword

- search for specific terms within an asset
- Example: processes with description 'medical reimbursement'

which are keyword

- search for assets with field values that have only two possible options
- Example: Business terms which are Critical Data Element

Containing keyword

- Use the containing keyword followed by an asset type to search for related assets.

Demo continued

Category	Example
Exact searches	"SSN"
Partial / Wild Card searches	"SOCIAL*"
Searches with Asset types qualified	report Customer*
Multiple search queries	Column Zip*, Column Addr* Data Elements Tax* , Data Elements Addr*
Connected Assets (AND)	(Column Addr*) in Table PARTY (Column Addr*) in Table Customer* ((Column Addr*) in Table Customer*) in schema HERMES
With key word	catalog sources with resource type 'Microsoft SQL Server' catalog sources with description
Containing	Domains containing business Term 'Postal*'
Which are	business terms which are critical data element
Not keyword	columns not related to data classification 'First Name'

Conclusion

- Cloud Data Governance and Catalog has the ability to model **complex enterprise technical data**.

- The **full value** is realized when Technical and Business Stewards can quicky **find** and **access** the data they need for **extending the model** and **resolving issues**. We hope the techniques introduced here will provide immediate help.

- Here are some KB reference articles to provide additional insight:

    - **Product Documentation for Searches:**
      https://onlinehelp.informatica.com/IICS/prod/DGC/en/index.htm#page/cloud-data-governance-and-catalog-working-with-assets/Search_and_browse_assets.html

    - **Master KB for searches:**
      https://knowledge.informatica.com/s/article/MASTER-KB-for-Search-Queries-in-Cloud-Data-Governance-and-Catalog-CDGC?language=en_US

Expert Panel

Q&A

Ratish Neelakandan
Solution Architect

Varun Sharma
Solution Architect