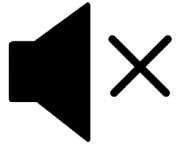

Scanning & Data Discovery Best Practices in Cloud Data Governance & Catalog

- Avanish Srivastava, Principal Solution Architect

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 **INFASupport YouTube channel** and **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.

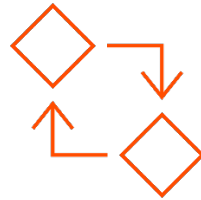
Feature Rich Success Portal



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>



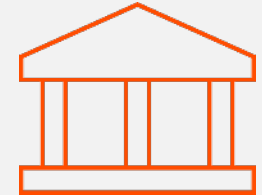
Communities & Support

<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

Introduction to
CDGC/IDMC

2

Metadata Scan
Overview

3

Scanner
Configuration
Process Flow

4

Data
Classifications

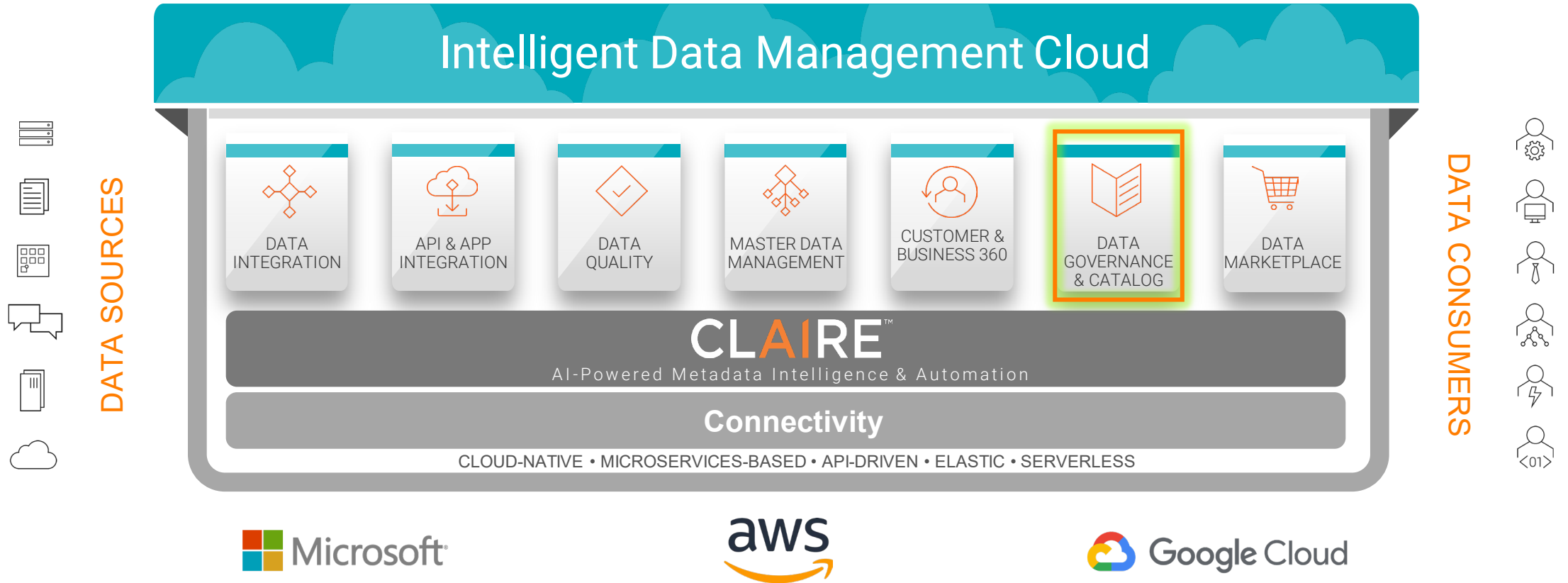
5

Demo

6

Q&A

Intelligent Data Management Cloud



Cloud Data Governance & Catalog Key Services

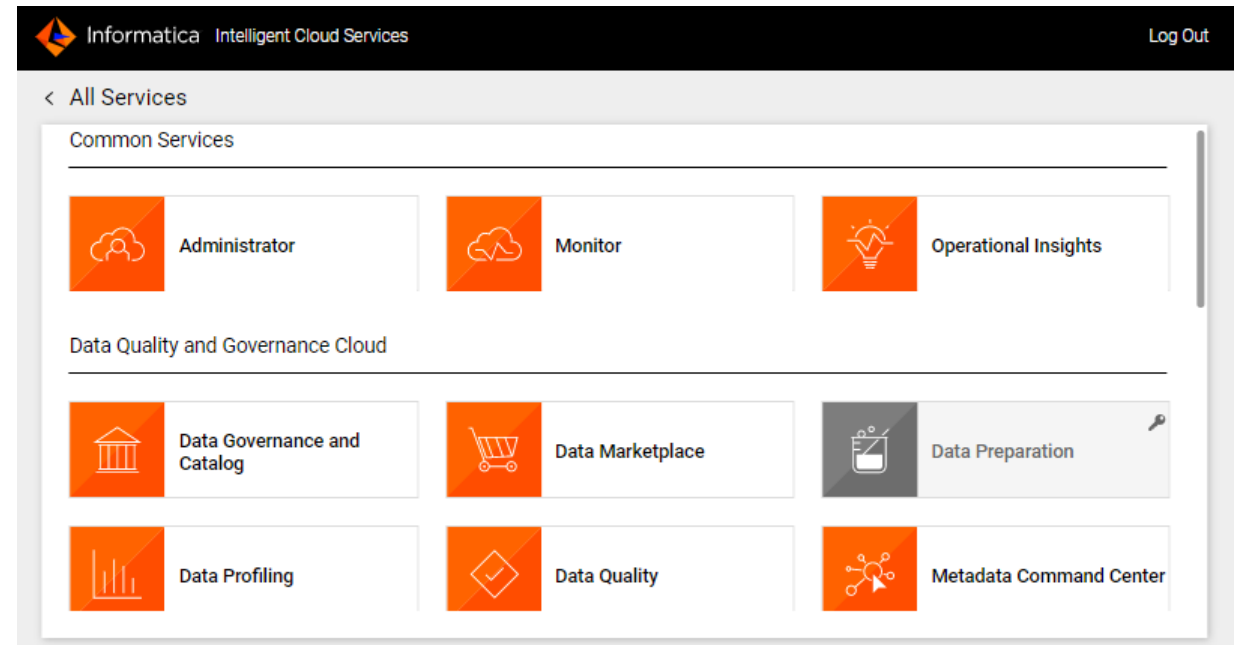
Administrator

Metadata Command Center

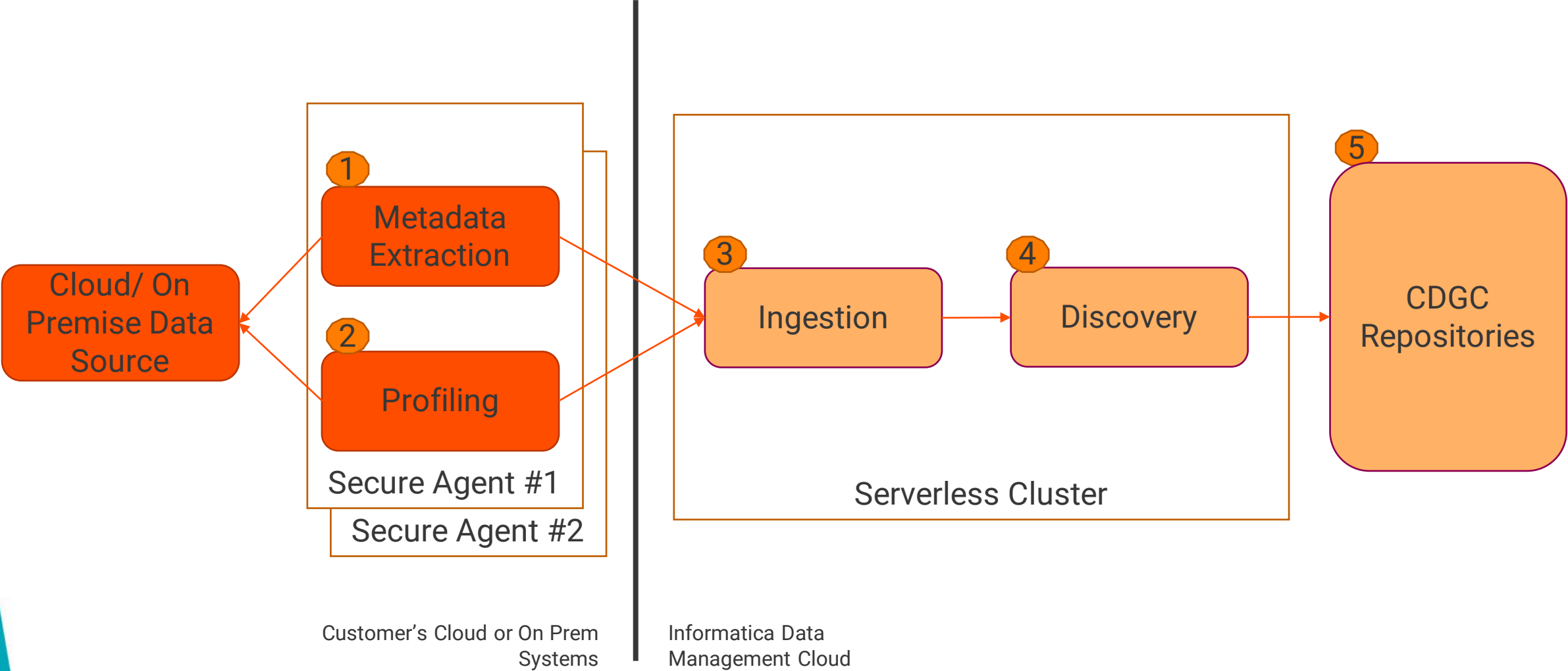
Data Governance & Catalog

Cloud Data Quality & Profiling

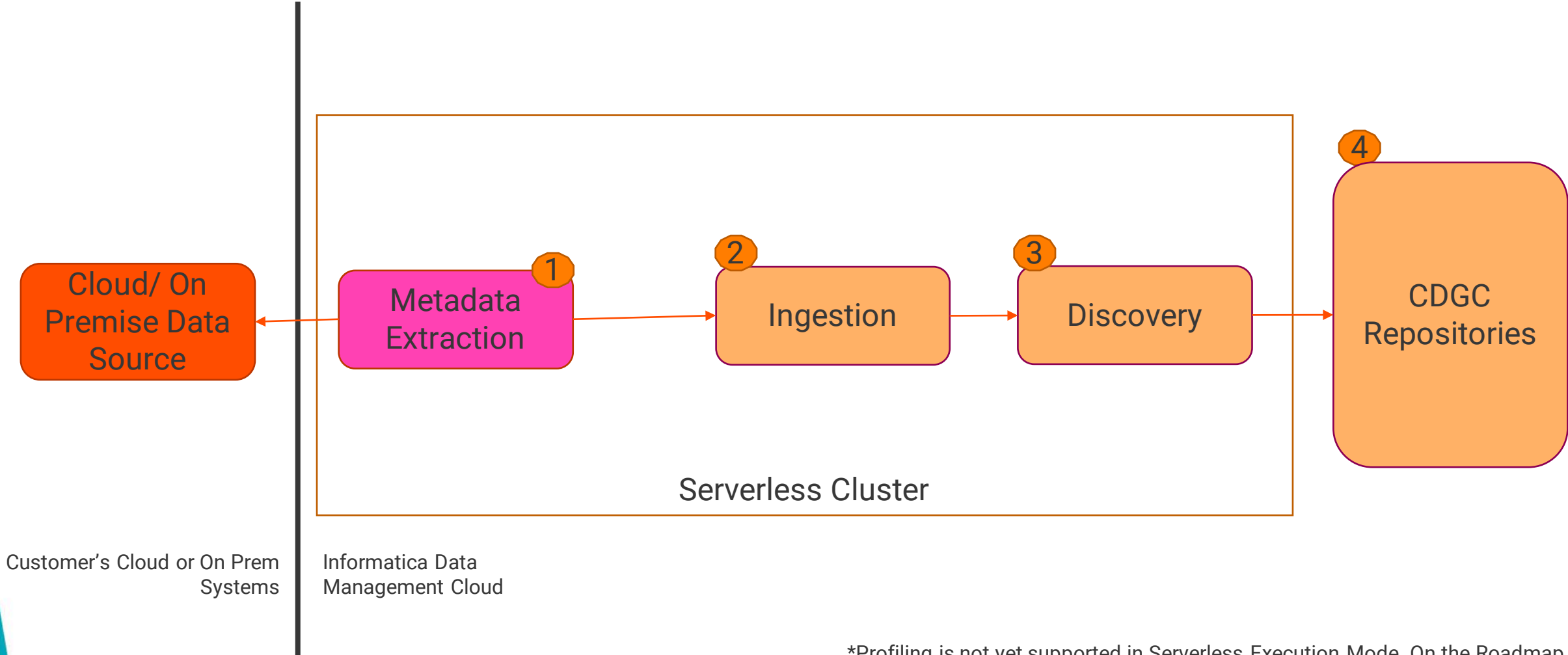
Data Marketplace



CDGC Secure Agent based execution



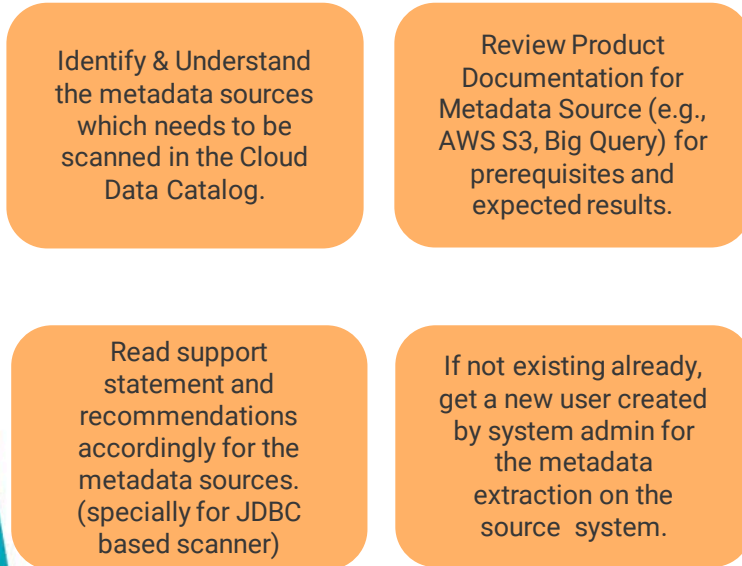
CDGC Serverless Execution



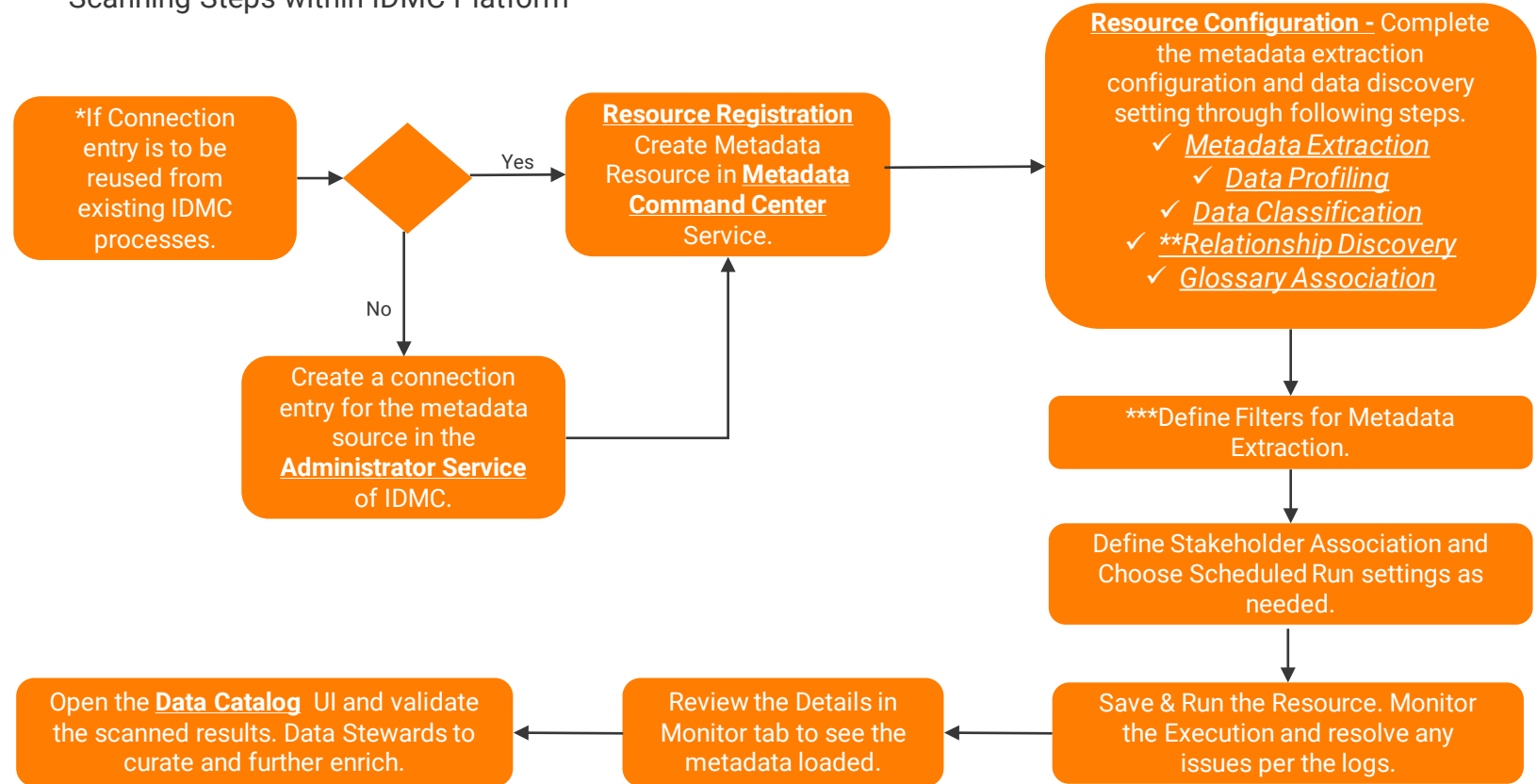
Profiling is not yet supported in Serverless Execution Mode. On the Roadmap.

Metadata Scan Process

Pework outside IDMC



Scanning Steps within IDMC Platform



* We recommend creating a new connection if possible due to different permissions needed for Metadata scan often.

** Relation Discovery may not be needed often. Read product doc.

*** We recommend using a metadata filter for better performance, cost optimization and avoid clutter in the catalog.

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.

Address Line 1

Save X

General Information >

Inclusion Rule Validate

Advanced You cannot toggle to Basic mode because you typed the expression in Advanced mode.

Attributes Find

Name	Data Type
Metadata	
COLUMN_COMMENT	String
NAME	String
PARENT_COMMENT	String
PARENT_NAME	String
Statistics	
AVERAGE	Numeric
FREQUENT_VALUES	Array
INFERRED_DATATYPE	String
MAX_VALUE	String
MIN_VALUE	String
NUM_PROFILED_VALUES	Numeric
STANDARD_DEVIATION	Numeric
VALUE_FREQUENCIES	Array(struct)

Select an item from the tree on the left to see the description. To see the list of suggestions in the editor, press Ctrl+Space in Windows or Cmd+Space in Mac. Type your expression below. Example: NAME LIKE "%Credit Cards" AND forall (FREQUENT_VALUES, v -> v RLIKE "4[0-9]12?[0-9]3[0-9]8")

```
(lower(name) like "%address%" or lower(name) like "%add?%1%" or lower(name) like "%addr?line1%") AND (size(filter(FREQUENT_VALUES, v -> trim(upper(v))) RLIKE $c1)) / size(FREQUENT_VALUES) >= 0.8f)
```

Address Data

General Information

Name: * Address Data

Description: This data entity classifies for Address Information.

Type: Entity

Inclusion Scope

Classifications: USA County x USA Zip x Street x Address Line 1 x USA City x Select

Include: All Any 3



References

Metadata Sources in CGDC

<https://docs.informatica.com/data-governance-and-quality-cloud/metadata-command-center/current-version/catalog-source-configuration/supported-source-systems.html>

Working with Assets in Catalog

<https://docs.informatica.com/data-governance-and-quality-cloud/cloud-data-governance-and-catalog/current-version/working-with-assets/preface.html>

Healthcare Case Study Webinar

<https://success.informatica.com/explore/tt-webinars/cloud-data-governance-healthcare-case-study.html>

Scanning SSIS, SSRS and MSSQL in CDGC

<https://success.informatica.com/explore/tt-webinars/configure-cdgc-scanners-with-microsoft-ssis-ssrs-and-sql-server.html>

Pre-Defined Data Classifications in CDGC

<https://docs.informatica.com/data-governance-and-quality-cloud/metadata-command-center/h2l/1715-predefined-data-element-classifications-in-cloud-data-governance/abstract.html>

Sizing Guidelines for Secure Agent in CDGC

<https://docs.informatica.com/data-governance-and-quality-cloud/metadata-command-center/h2l/1651-sizing-recommendation-and-performance-optimization-for-meta/abstract.html>

DEMO

Q&A