

05.14.2020

Data Governance - Data Lake

Webinar Series

Asit Rath, Manager, IPS

Ratish Neelakandan, Solution Architect, IPS



Informatica™

Agenda

1

Use case
Decomposition/prioritization

2

Pilot Implementation flow

3

Persona Identification and
Org Model

4

Determine 'Day in the life of'
Scenarios

5

Axon, EDC and DQ
Operating Model

6

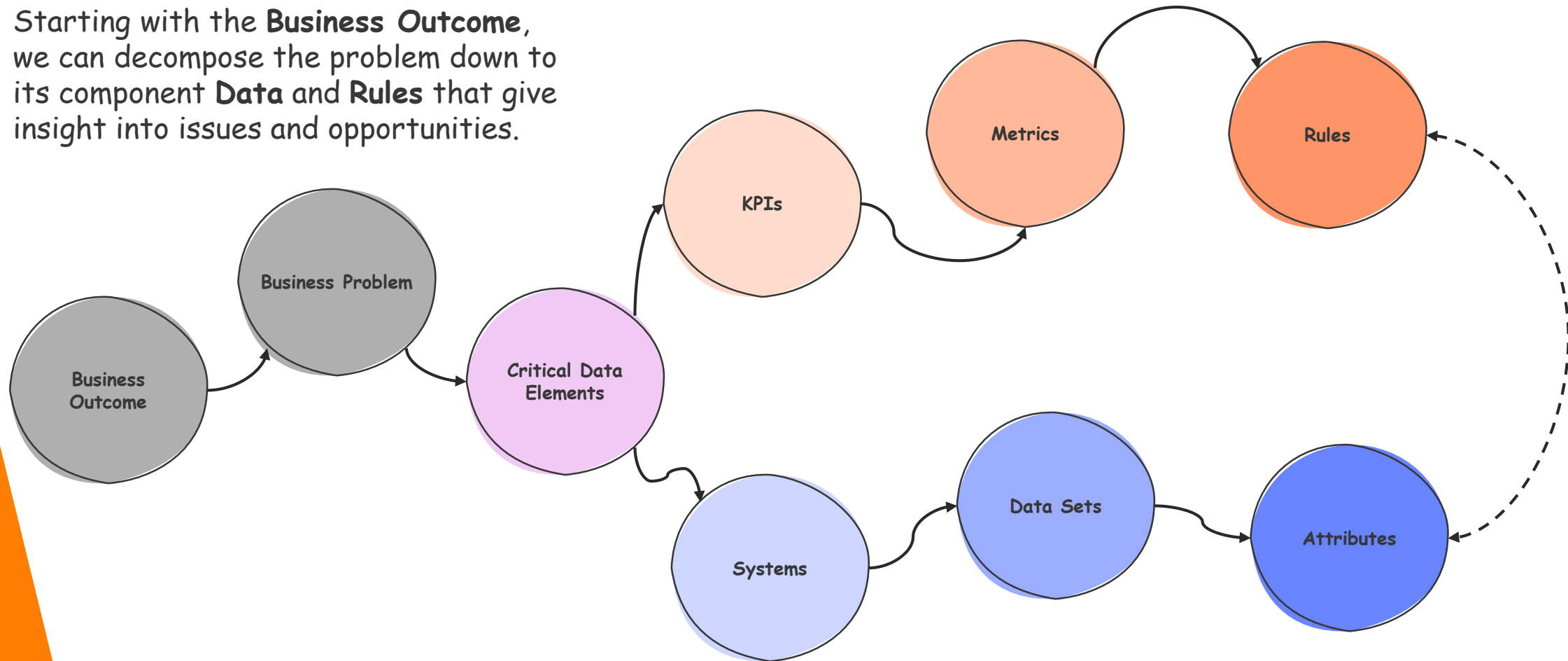
'Day in the life of' execution
flows

7

Short Demo

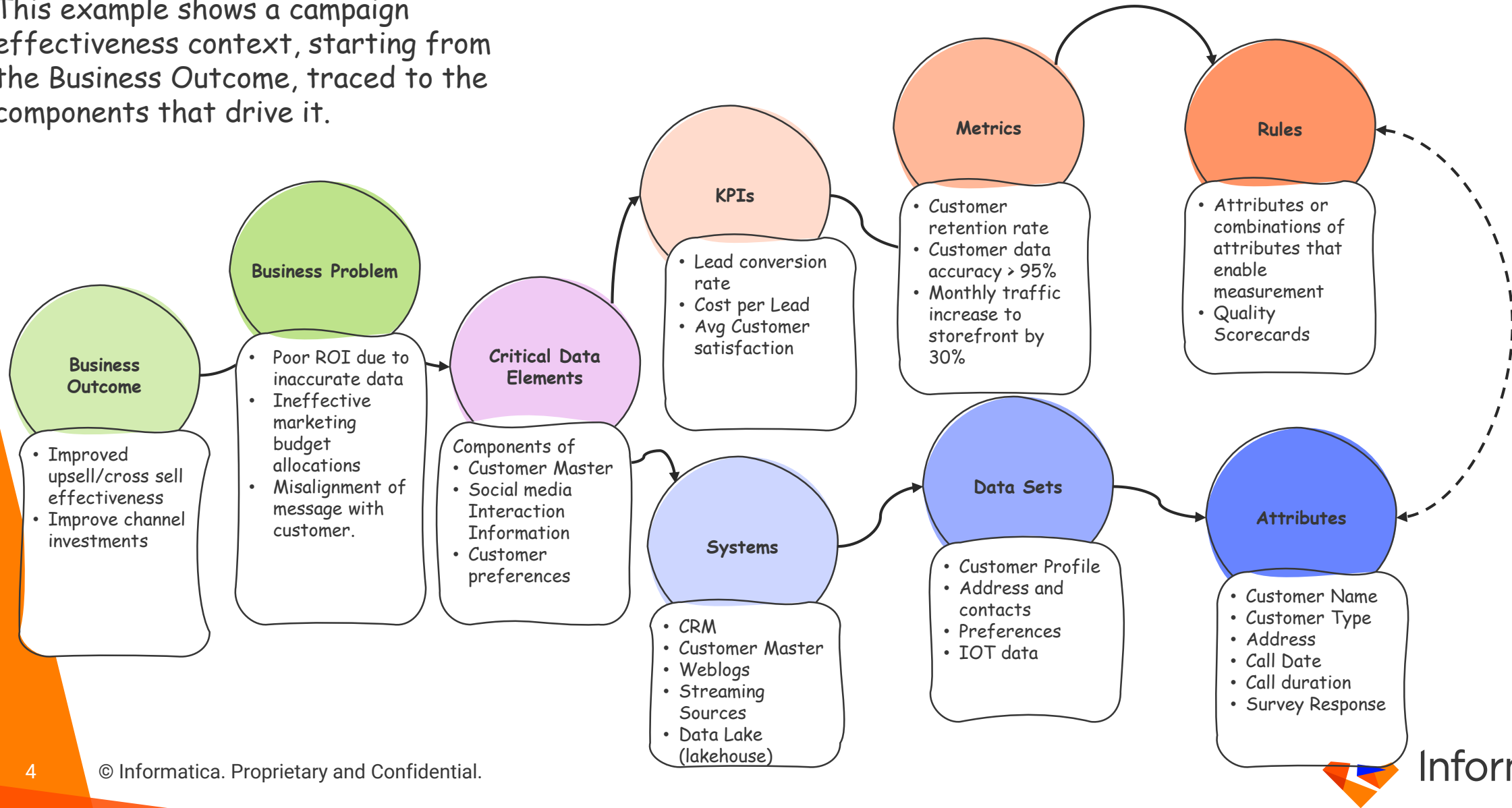
Data Governance: Business Problem Decomposition

Starting with the **Business Outcome**, we can decompose the problem down to its component **Data** and **Rules** that give insight into issues and opportunities.



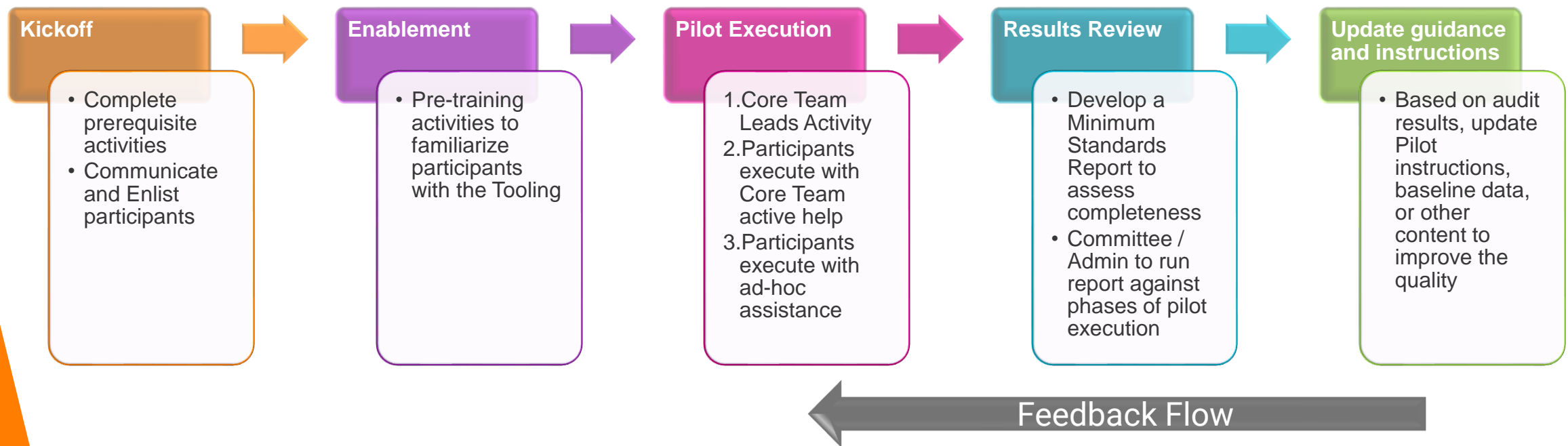
Decomposing a Marketing Effectiveness Use Case

This example shows a campaign effectiveness context, starting from the Business Outcome, traced to the components that drive it.



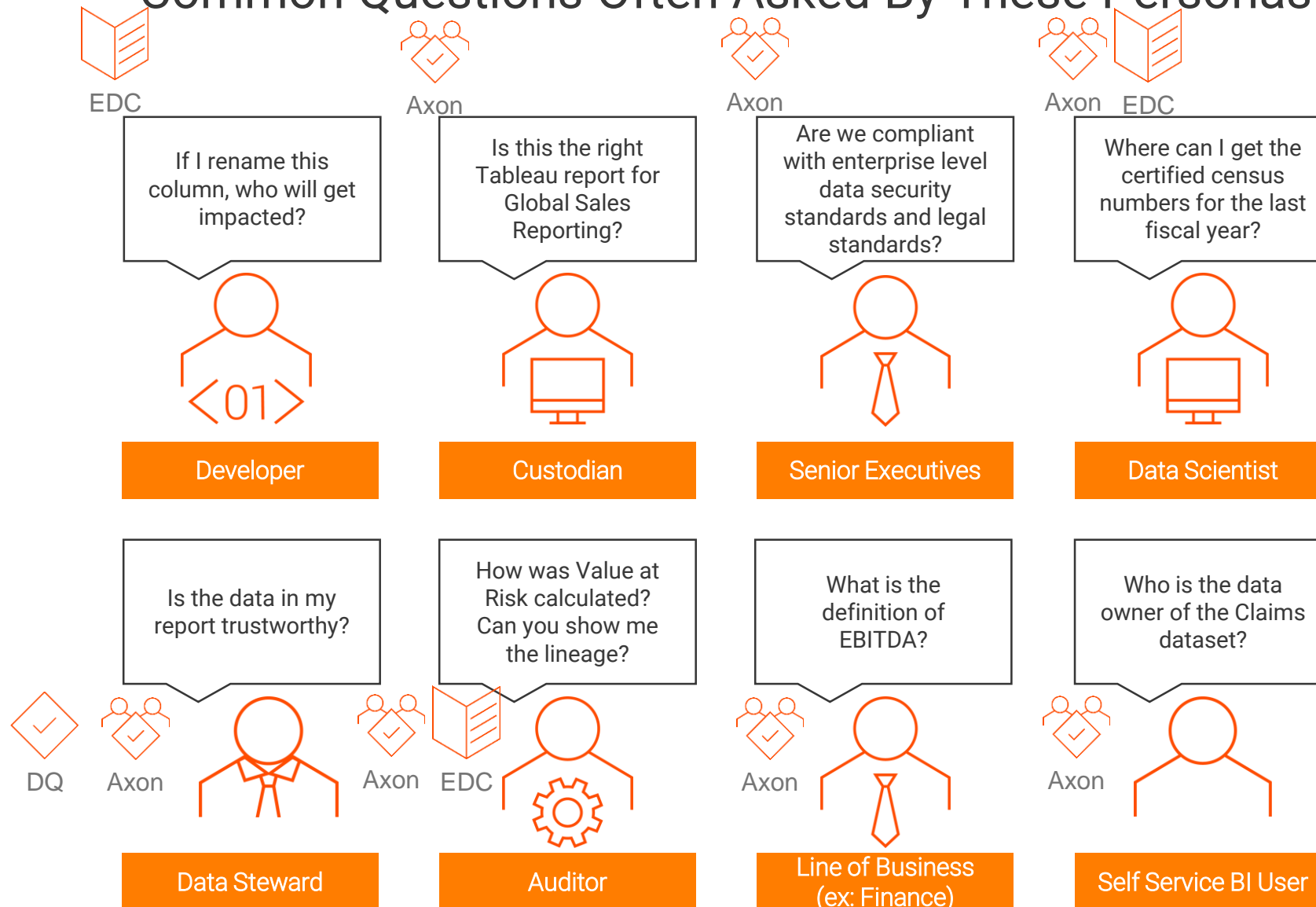
Pilot Implementation Process Flow

The following flow is essential to managing the inclusion of feedback into the pilot flow, for later adopters and the development of reusable processes.

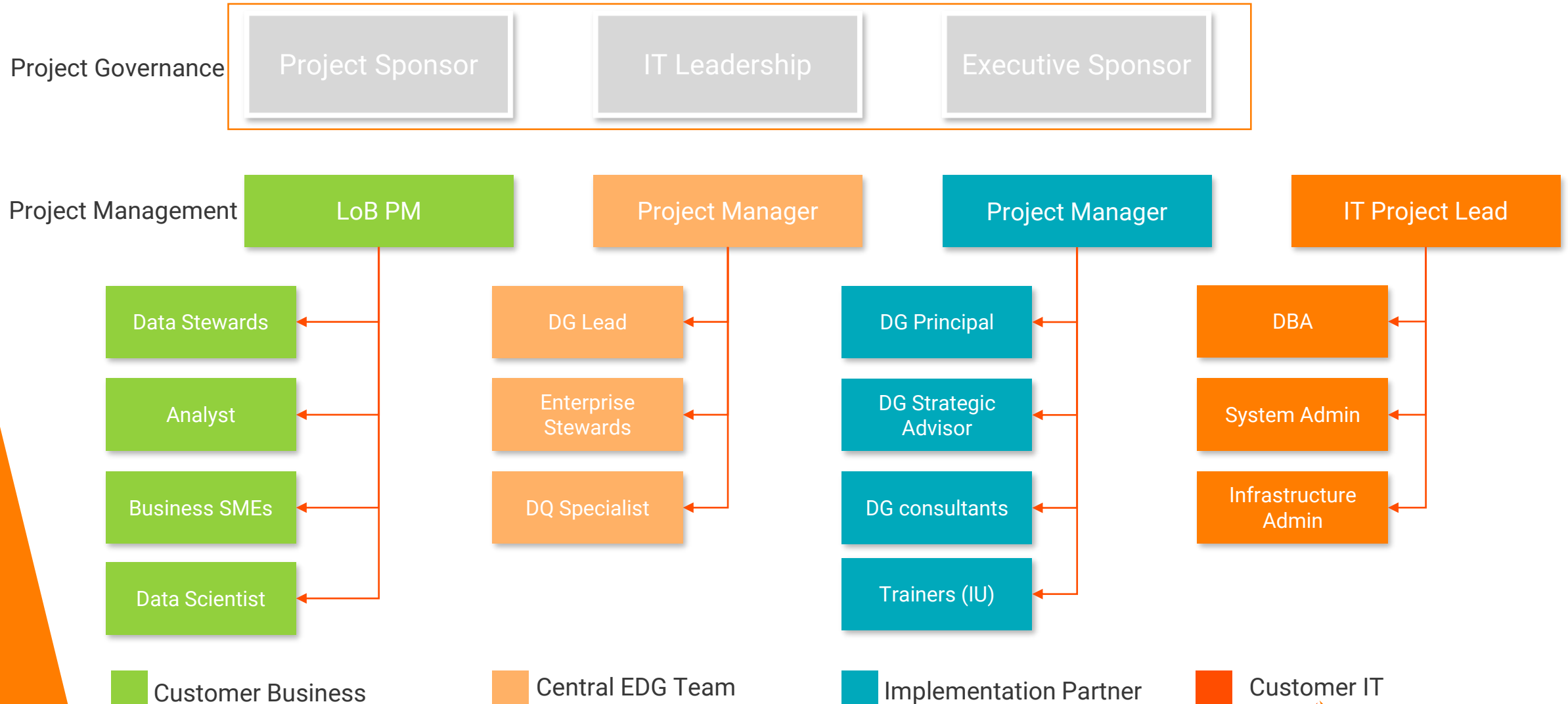


Program Overview & Informatica Tools

Common Questions Often Asked By These Personas



Typical DG Engagement Organization



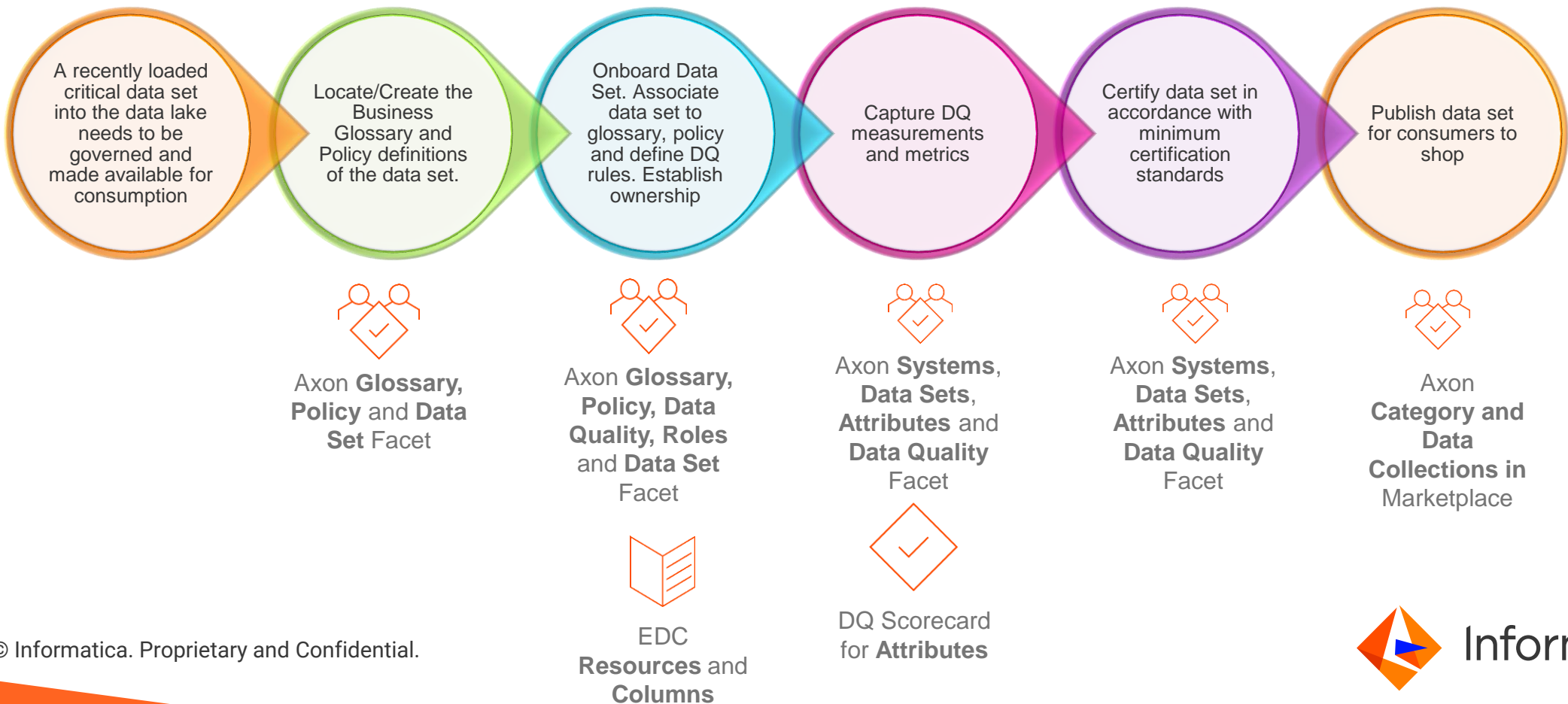
Publishing Critical Data Sets to the Storefront

How do I make my critical data sets in the Lake available for consumers and track access and usage



The Data Steward publishes a data asset after certifying it.

*This is a **Publisher** Use Case*

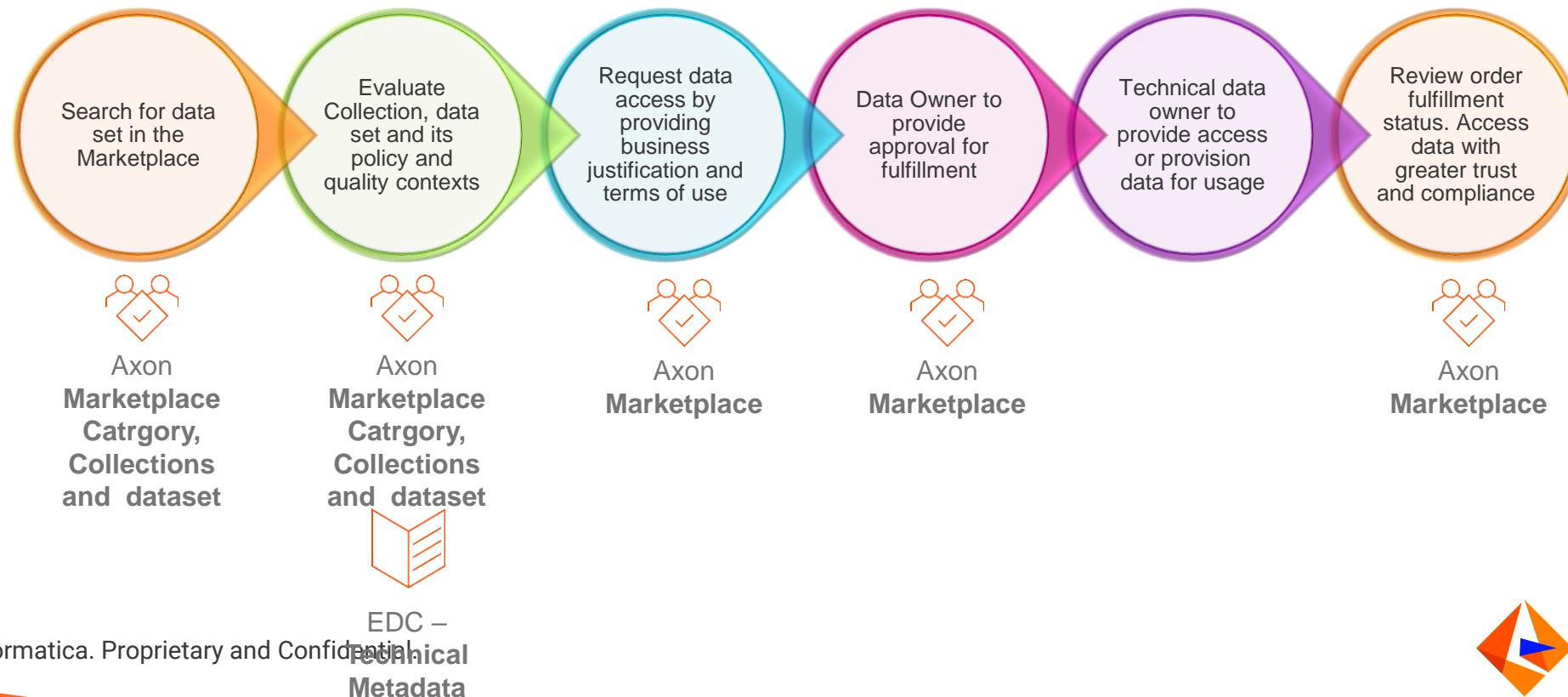


Shopping for trusted data in the Lake

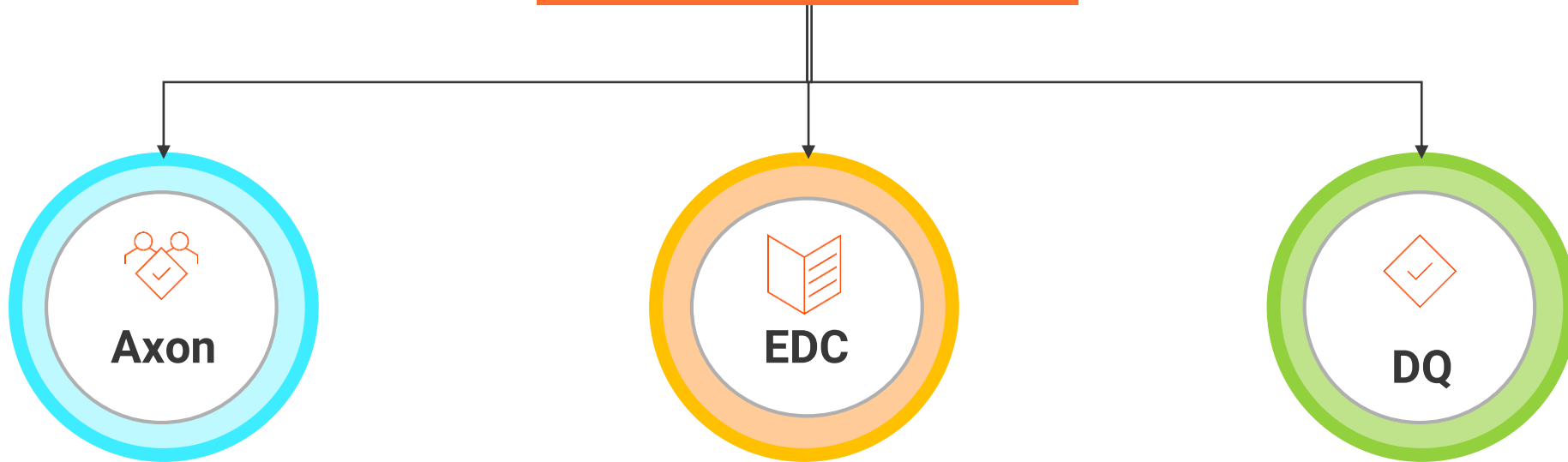
The Data Scientist/Self-service consumers to easily search, understand and access the right data needed for analytics

*This is a **Consumer** Use Case*

How do I shop for trusted data in the Lake for my Analytics



Operating Model



Axon Facets

Glossary Taxonomy

Roles and privileges matrix

Custom Fields

Drop Down Customization

Reusable Workflows

Categories

Data Delivery option

Security Model

Roles and Privileges
Matrix

Custom Tags

Domain Rules

Security Model

Roles and
Privileges Matrix

Metrics Groups

EDC Security Model - Sample

Groups	Administrator	Catalog Administrator/local administrators	Stewards/Analysts/Data Owners	All Read Only Users	Developers(API Users)
Sub Groups					
Roles					
Privilege Name	Production	Production	Production	Production	DEV/TST
Edit Reference Table Metadata	X	X			
Catalog Service					
API Privilege					
REST API privilege		X	X		X
REST API User privilege		X	X		X
Catalog Privileges					
Catalog View	X	X	X	X	X
Catalog Edit		X	X		X
Application Configuration	X	X	X		
Domain Creation		X			
Domain Curation			X		
Resource Management					
Admin - View Resource	X	X	X	X	X
Admin - Edit Profiling	X	X	X		
Admin - Edit Resource	X	X	X		X
Domain Management					
Admin - View Domain and Domain group	X	X	X	X	X
Admin - Edit Domain and Domain group		X	X		
Data Privileges					
View Data			X		
View Sensitive Data			X		
Admin - Create Attribute		X			X
Admin - Monitoring	X	X	X		X

Roles

Role Privileges

Map roles to AD groups/users

Map Groups to EDC resources

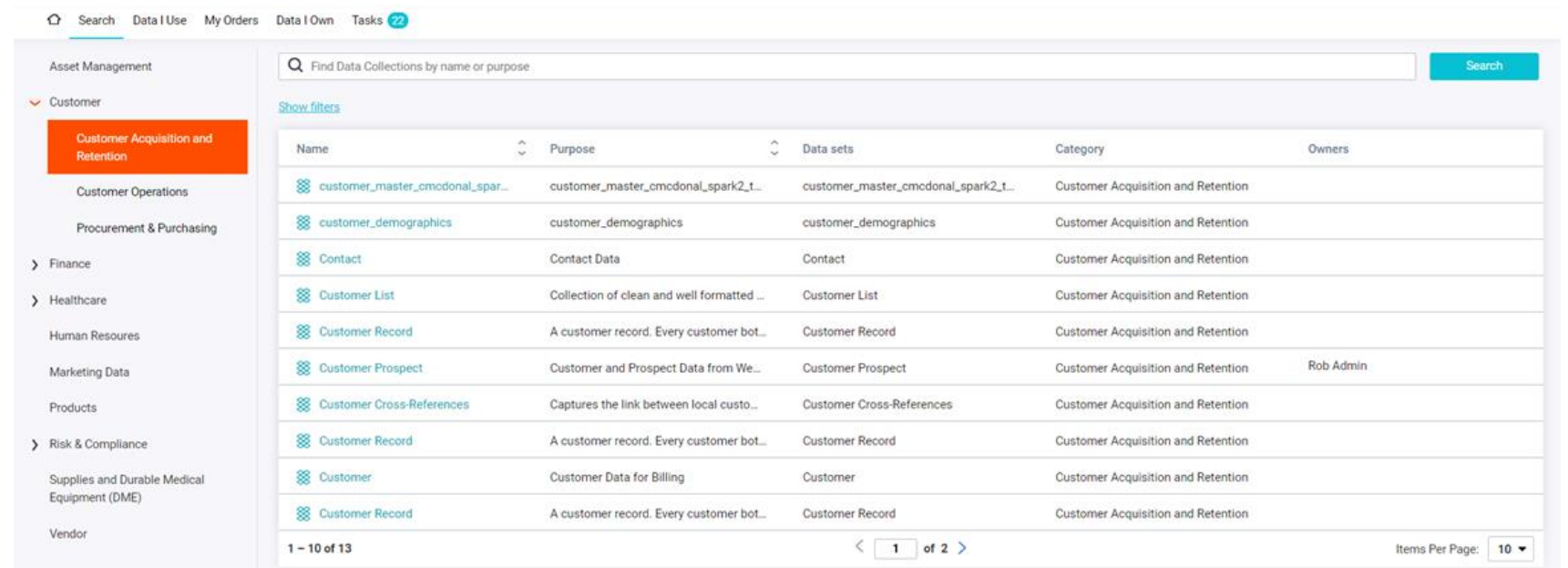
Operating Model – Marketplace Categorization (Sample)

Categories

Data Delivery option

Sub Categories

Map Data Sets to
Categories/Sub
Categories



The screenshot displays the Informatica Data Marketplace interface. On the left, a navigation pane shows a hierarchy: Asset Management > Customer > Customer Acquisition and Retention (highlighted in orange). The main area features a search bar and a table of data collections. The table has columns for Name, Purpose, Data sets, Category, and Owners. The 'Category' column for all items is 'Customer Acquisition and Retention'. The 'Owners' column shows 'Rob Admin' for the 'Customer Prospect' item. The table lists 13 items, with the first 10 visible on this page. A pagination bar at the bottom indicates '1 - 10 of 13' items and 'Items Per Page: 10'.

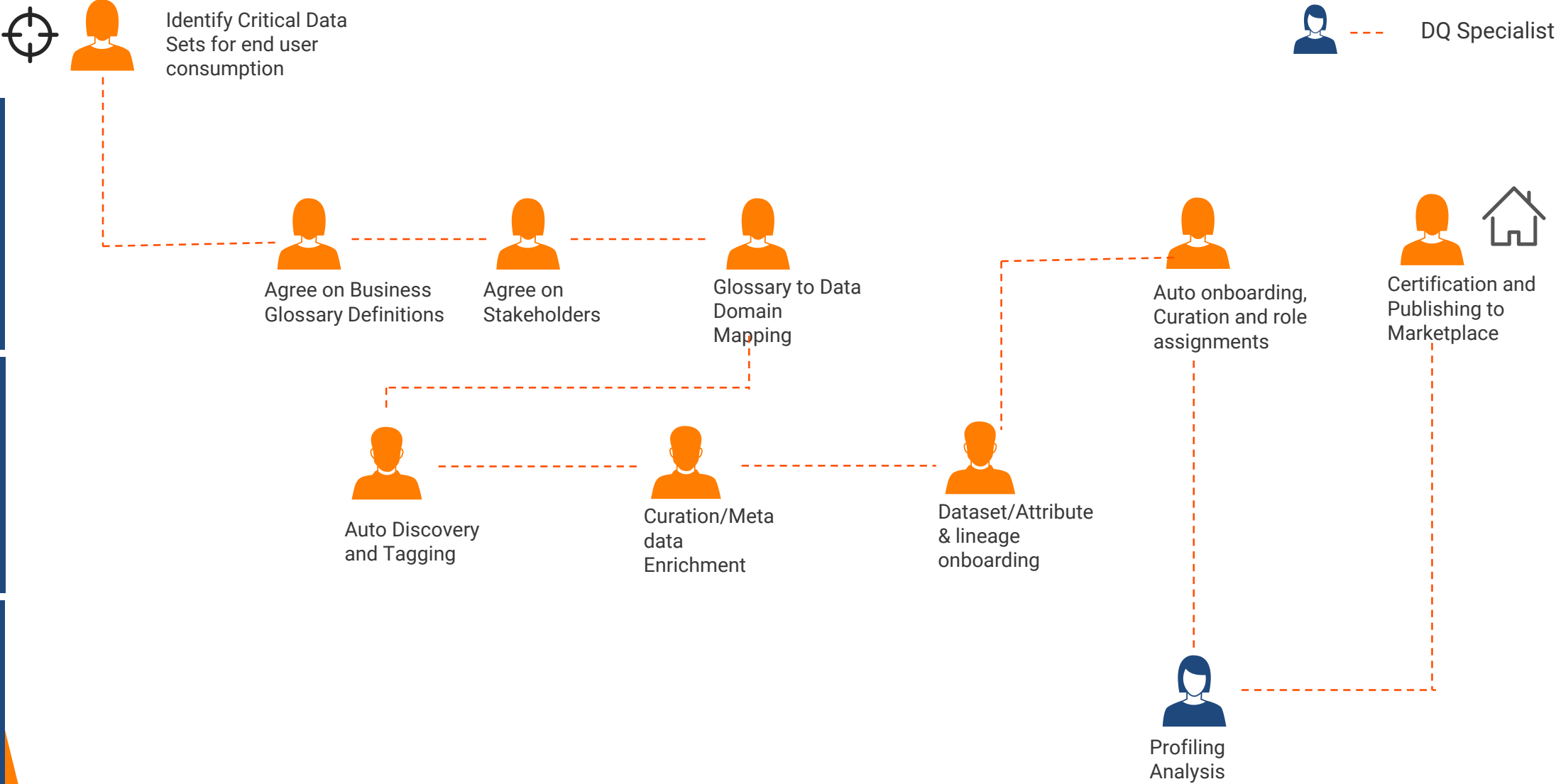
Name	Purpose	Data sets	Category	Owners
customer_master_cmcdonal_spar...	customer_master_cmcdonal_spark2_t...	customer_master_cmcdonal_spark2_t...	Customer Acquisition and Retention	
customer_demographics	customer_demographics	customer_demographics	Customer Acquisition and Retention	
Contact	Contact Data	Contact	Customer Acquisition and Retention	
Customer List	Collection of clean and well formatted ...	Customer List	Customer Acquisition and Retention	
Customer Record	A customer record. Every customer bot...	Customer Record	Customer Acquisition and Retention	
Customer Prospect	Customer and Prospect Data from We...	Customer Prospect	Customer Acquisition and Retention	Rob Admin
Customer Cross-References	Captures the link between local custo...	Customer Cross-References	Customer Acquisition and Retention	
Customer Record	A customer record. Every customer bot...	Customer Record	Customer Acquisition and Retention	
Customer	Customer Data for Billing	Customer	Customer Acquisition and Retention	
Customer Record	A customer record. Every customer bot...	Customer Record	Customer Acquisition and Retention	

Data Lake - Data Provisioning - Context

- ❑ Key concepts for a successful Data Provisioning and Order fulfillment:
 - ✓ Establishing processes to reach consensus on publishing critical Data Lake data sets
 - ✓ Policy and Usage based consumption of data from the Data Lake

- ❑ The following execution flows will help in understanding:
 - ✓ A common approach for identifying critical data sets, onboard to the governance platform and publish them
 - ✓ Operationalizing the order management process for data consumption




Critical Data Set Onboarding Process Flow



Order Fulfillment – Consumer Execution Flow

Axon/marketplace

EDC

-  Data Consumer
-  Data Owner
-  Technical Data Owner



Data Consumer initiates a data access order



Searches data sets in Marketplace



Accepts terms of usage and requests data access



Reviews and accepts data access request



Reviews and provides data access/provisions data



Operationalize Order Management process



Evaluates technical assets

Execution Flow Steps- Operational Runbook



Outcome



How-To instructions



Do's and Don'ts



Troubleshooting



Get Help



Data Steward

Curation/Meta
data
Enrichment

Short Demo

We're Ready to Help!

Different ways we can help

- Follow up Sessions
- Technical and Business Workshops
- Assessments
- Customized Engagements
- Technical and Business Advisors
- Implementation Support

For follow up and additional questions, please reach out to:

Jaideep Singh
(Jsingh@informatica.com)

**DG & Privacy Journey
Lead- Professional
Services**

Chris Main
(Cmain@informatica.com)

**Sr. Director-Professional
Services**

Paul Yoo
(Pyoo@informatica.com)

**Sr. Director-Professional
Services**

Jason Beard
(jbeard@informatica.com)

**Sr. Director-Data Strategy
& Governance**



Informatica
Professional Services

Q&A