

Jan 14, 2019

# Intelligent Data Governance

## Using Axon, EDC, IDQ

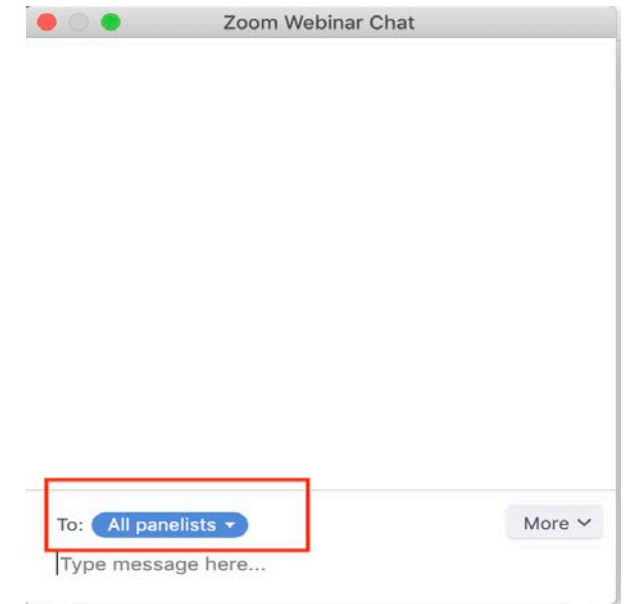
Srinivasa Gopal

*Principal Technologist – Customer Success*



Informatica™

# Housekeeping Tips



- Today's Webinar is scheduled to last **1 hour including Q&A**
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the **Chat option** and we will respond at the end of the presentation
- The webinar is **being recorded** and will be available to view on our **INFASupport YouTube channel**. The link will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

<https://success.informatica.com>



Bootstrap product  
trial experience



Enriched Onboarding  
experience



Product Learning  
Paths and weekly  
Expert sessions



Informatica  
Concierge with  
Chatbot integrations



Tailored training and  
content  
recommendations

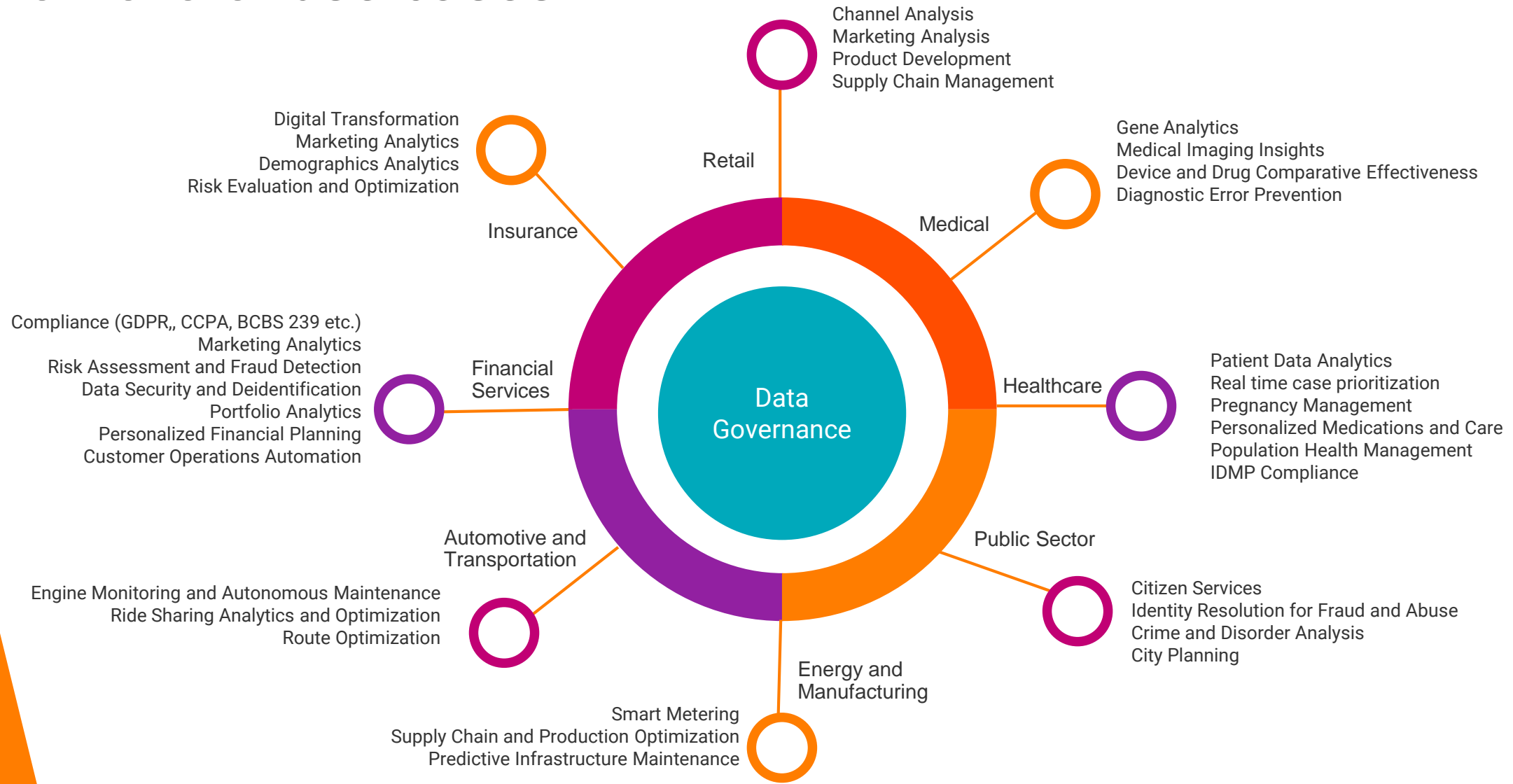
# Safe Harbor

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

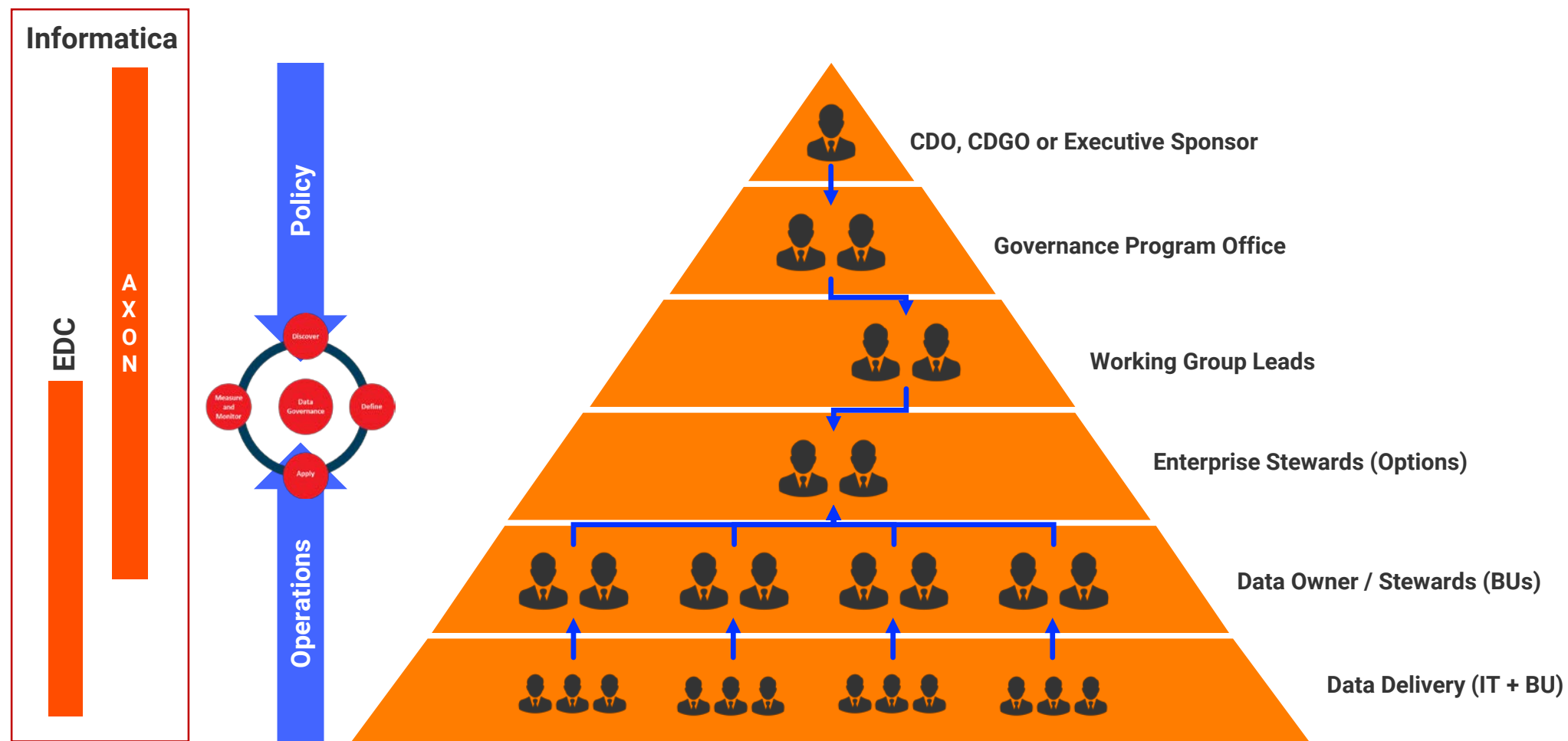
- UseCases & Overview of Axon – Data Governance
- Automatic Data Discovery, Cataloging using EDC
- Automate Fit for purpose Data Quality using IDQ
- Demo
- Q&A

# Portfolio of use cases

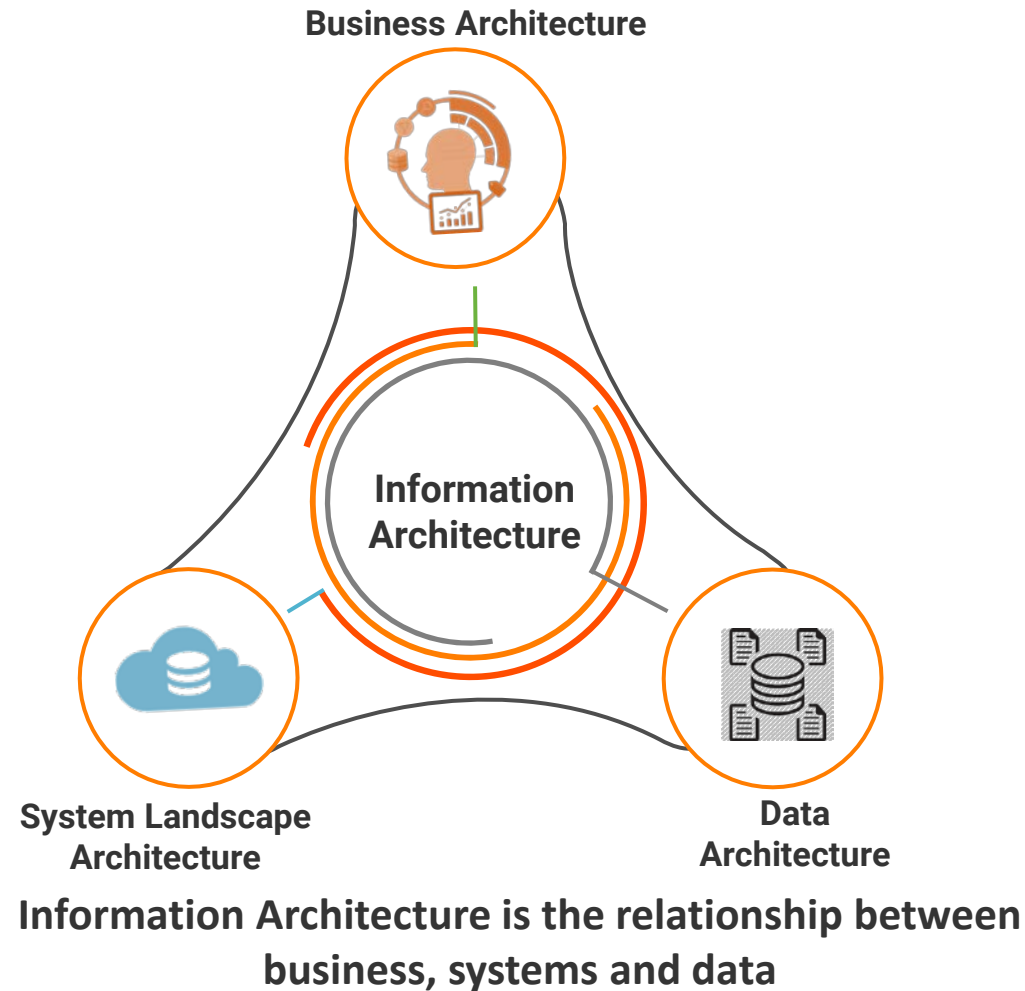


# Business Problem

Alignment across Policy and Organization tiers



# Connected Data Governance – Information Architecture



## Business Architecture

Understanding the business need

*Business groups, processes, controls, connections*

## System Landscape Architecture

Understanding how applications enable business processes, record and transport information

*Systems, Interfaces*

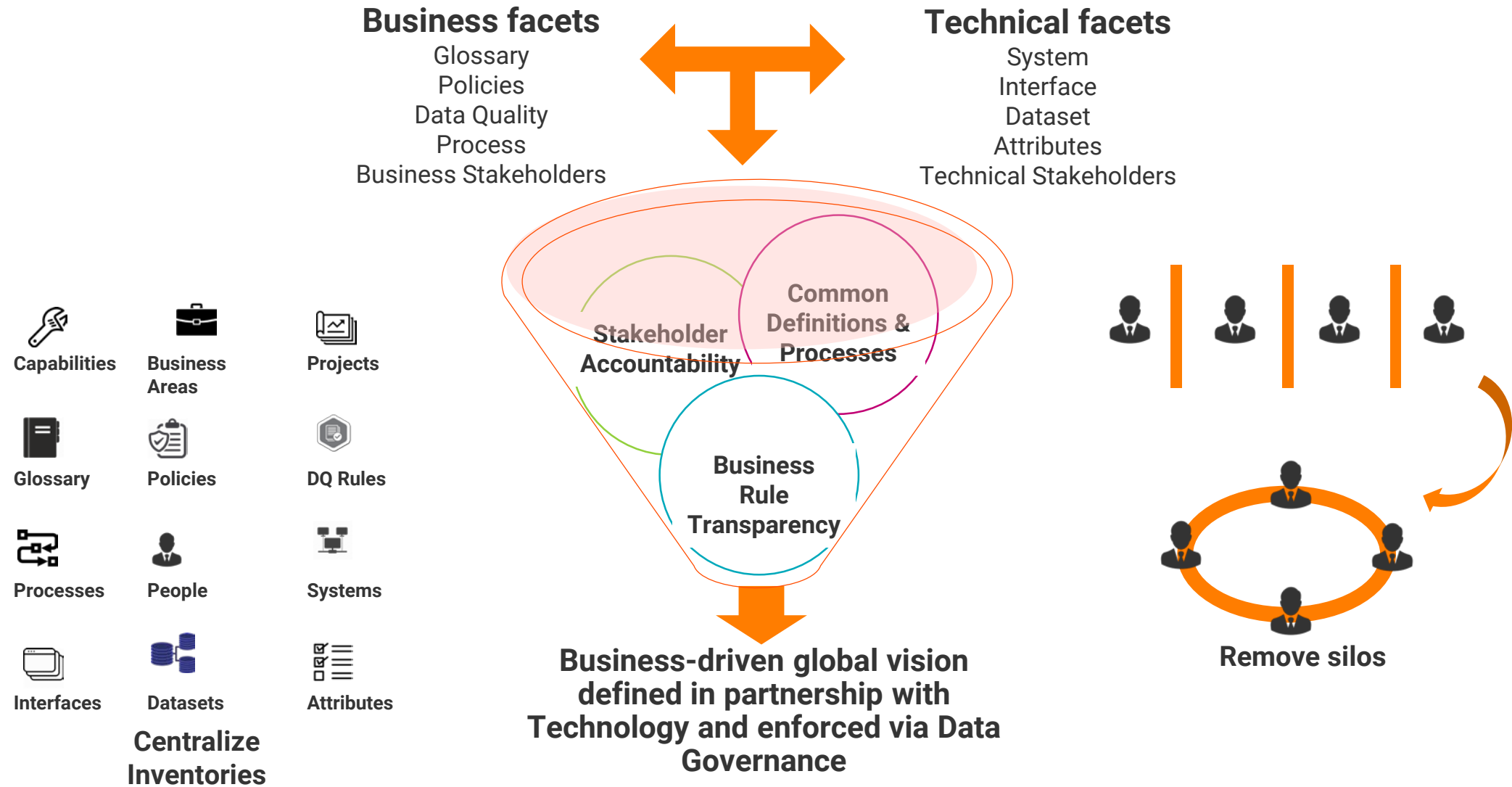
## Data Architecture

Understanding our information and how it binds business and IT together

*Data, definitions, standards, quality stewardship*



# Axon EDG Management Concept



# Automate Data Discovery, Cataloging, and Context

Leverage ML and AI to find critical data across structured and unstructured sources

Onboard discovered data automatically with oversight and control

Automatically tag data with business context to help users assess relevance

The screenshot displays the Informatica Data Catalog interface, organized into three main sections:

- Top Section: Data Discovery**
  - Filters: Location, Customer, Person.
  - Buttons: Show All.
  - Table with columns: Name, Business Title, Data Domains, Null | Distinct | Non-Distinct %, Source Data Type, Inferred Data Types.
  - Rows: cust\_email (Email), cust\_firstname (FirstName), cust\_gender (Gender), cust\_id (CMD-ID).
- Middle Section: DATA ATTRIBUTE INSTANCES**
  - Review Status: Accepted, CLAIRE Discovered.
  - Table with columns: Ref, Name, Definition, Review Status, Field Name, Confidence Score (%), Data Set.
  - Rows: Attribute Name 01 to 33, with some marked as CLAIRE Discovered.
- Bottom Section: Customer (customer\_master\_CS)**
  - Overview tab selected.
  - Description: A party that receives or consumes products (goods or services) and has the ability to choose between different products and suppliers.
  - Sample Columns table with columns: Name, Business Title, Data Domains.
  - People sidebar: Data Owner (Gaurav Pathak), Data Steward (Khan Tan), Experts (Abhiram Gujjuwar).

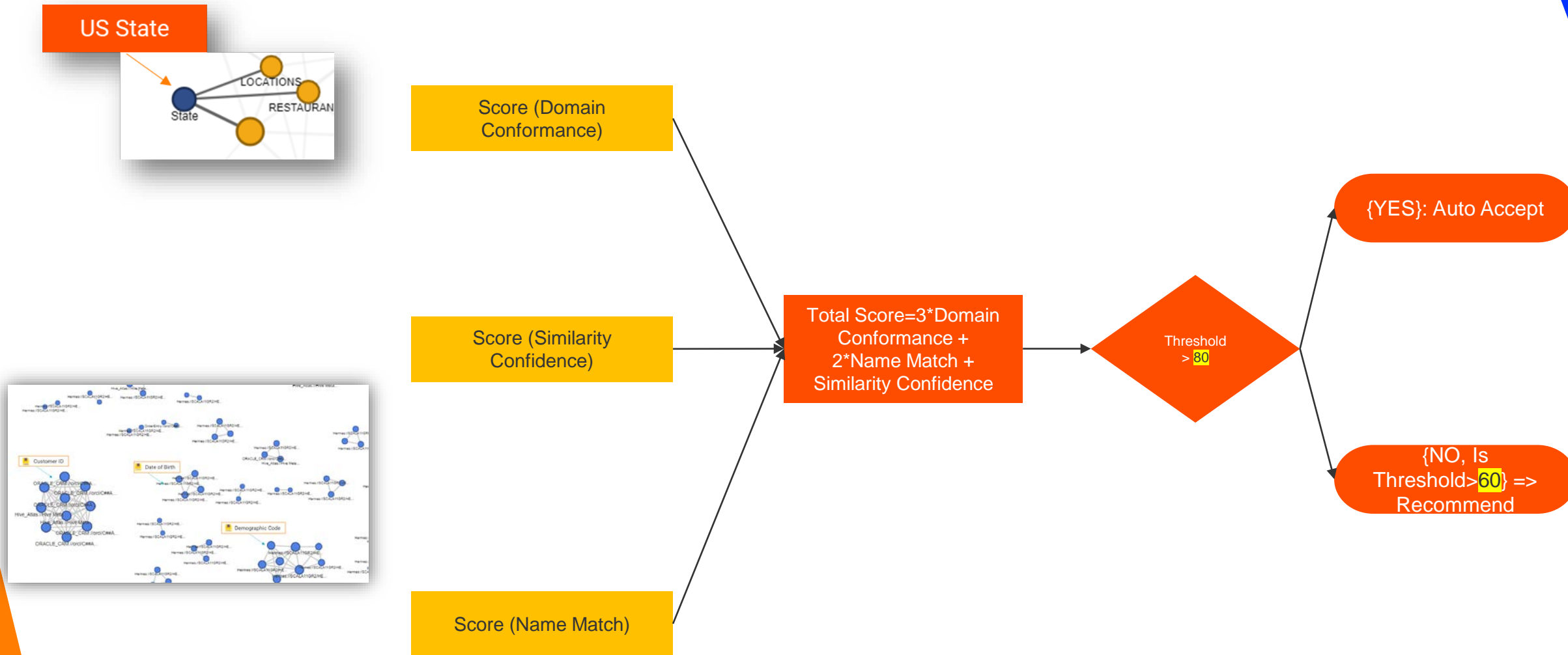
# Intelligent Glossary Association

## CLAIRE Match

Column Name	Business Name (Actual Auto Associations done by the system)
FIN_OPS_CNTCT_CD	Financial Operations Contact Code
IC_RATE_PCT	Indirect Cost Rate Percent
AR_ITEM	Accounts Receivable Item
AR_ITEM_ACT_ENTRY_AMT	Accounts Receivable Item Activity Entry Amount
EXTRACT_DT	Extract Date
BILL_ACCT_RECON_FLG	Bill Account Reconciliation Flag
PYRLL_TRANS_UNALLW_IND	Payroll Transaction Unallowed Indicator
CLASS_EFFDT	Class Effective Date
CLM_SOR_CD	Claim System of Record Code

- Match English phrases with technical names using sequence alignment
- **Sequence Alignment / Delete-only Edit Distance** : The business term names that align well with asset names are sought. This approach can capture obvious abbreviations of business terms.
  - HEALTH PROGRAM CONSULTATION (Business Term Title)
  - H- - LTH P - - G - - MC- NS - LT- T- -N (Asset name)
- **Synonym dictionary** : If available, user provides a dictionary of commonly used synonyms/abbreviations in technical asset names within the organization. This dictionary is used to improve glossary matching
- Additionally, prefix ignore options for discarding common technical prefixes(like TBL, VW etc for better matches)
- Name match is in tech preview in 10.2.2 and will not be available to customers by default

# Intelligent Business Glossary Associations



# Business Glossary Assignment Report



## Data Steward

Understand glossary association coverage for a resource in one place. Curate recommendations directly from the report

Assets with unassigned Business Titles names

432 Assets

Asset Name	Asset Type	Business Title	Recommended Business Title
cust_23 SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	72 Customer ✓ x 54 Employee ✓ x 54 Sales ✓ x
sales_data SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	90 Sales ✓ x 65 Sales Cost ✓ x 54 Custoer ✓ x
store_inventoruy SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	81 HR ✓ x 72 Actual Cost ✓ x 65 Finance ✓ x
card_sales_data_21 SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	83 Flight ✓ x 70 Actual Cost ✓ x 64 Card ✓ x
HR_confid_data SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	72 Delay ✓ x
flight_delay SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	92 HR ✓ x
raw_cust_analysis_GDP SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	57 Analysis ✓ x
mailing_list_NAM SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	69 Cost ✓ x 72 Actual Cost ✓ x 64 Mailing ✓ x
cust_data SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	
emp_data SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	89 Employee ✓ x 73 Actual Cost ✓ x 62 Data ✓ x
spending_patrn_analysis SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	80 Sales ✓ x 55 Pattern ✓ x 51 Finance ✓ x
risk_score SAPHana11:/H11/REP_ILMN_ECC/DD0	Table	Add Business Title	90 Risk ✓ x 90 Score Cost ✓ x

1-20 of 100 < 1 of 5 > Items per page: 20

- Users can go to the Business Term Assignment Report at the resource level to look at all business term recommendations in bulk and curating them.
- Includes filters for asset names, associated items and recommended terms.



# Automate Fit for Purpose Data Quality

Automate rule generation from technical descriptions of standard measures and metadata

Automate rule enforcement in systems and business processes

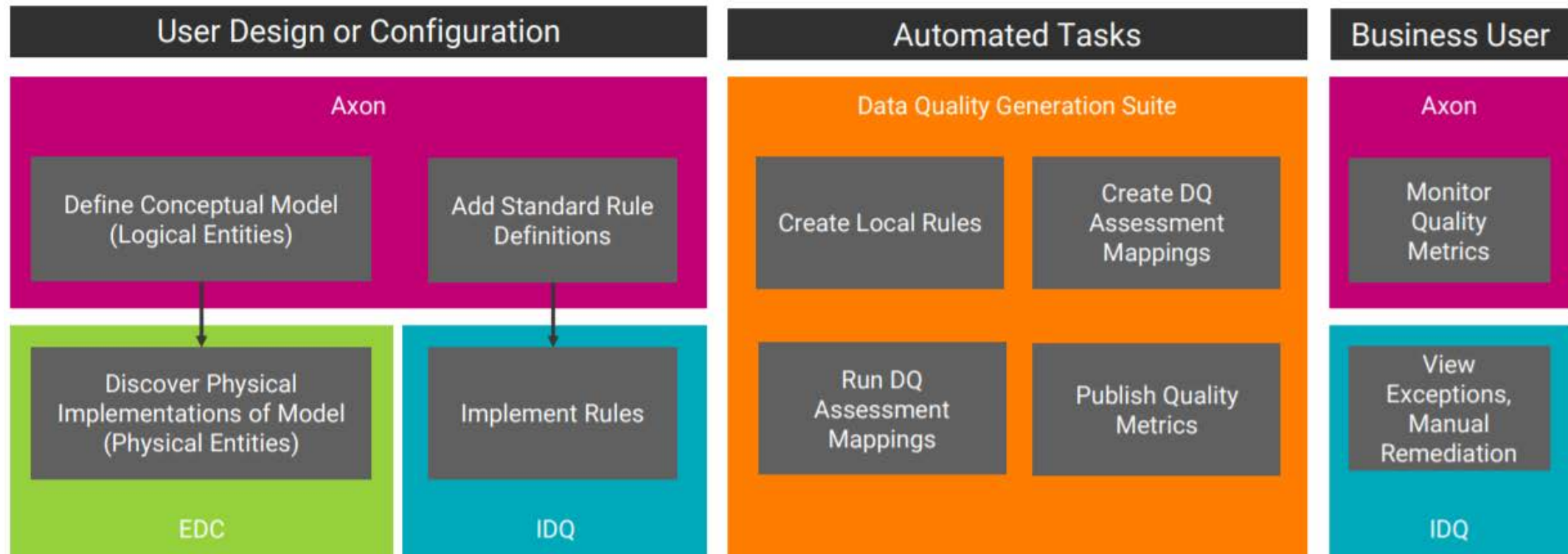
Monitor quality improvement across systems and processes over time

LOCAL DATA QUALITY RULES

Standard Ref.	Local Ref.	Type	Description	Target	Threshold	Last Result	System	Axon Status
31	DORULE-31	Accuracy	This is DQ Rule number 31	100%	80%	64%	FXM	Active
STRAT-2	DQ-2	Validity	CMD IDs should be of length 8 and leading zeroes are not accepted.	90%	98%	100%	CMD	Active
1	DQ-1	Completeness	All active Legal Entities should have a CMD ID.	95%	90%	96.87%	CMD	Active
STRAT-2	DORULE-3	Validity	CMD IDs should be of length 8 and numeric.	99%	98%	96%	FXM	Active
STRAT-2	DQ-10-1	Validity	Only numeric id with no leading zero's are valid	98%	96%	91.32%	KYC	Active
	DQ-35	Completeness	Check for CMD ID is not null	90%	84%	94%	CMD	Active

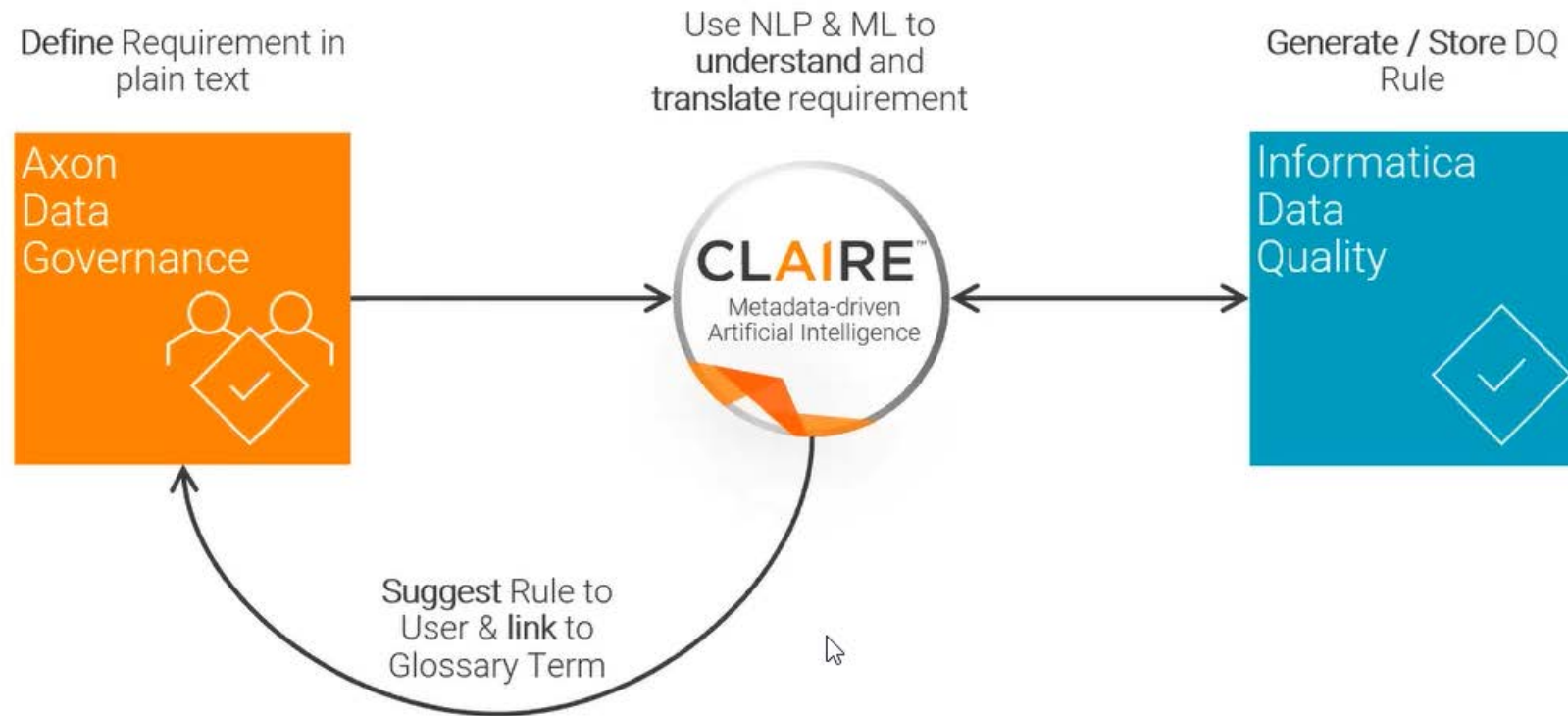


# Automating Data Quality for Data Governance



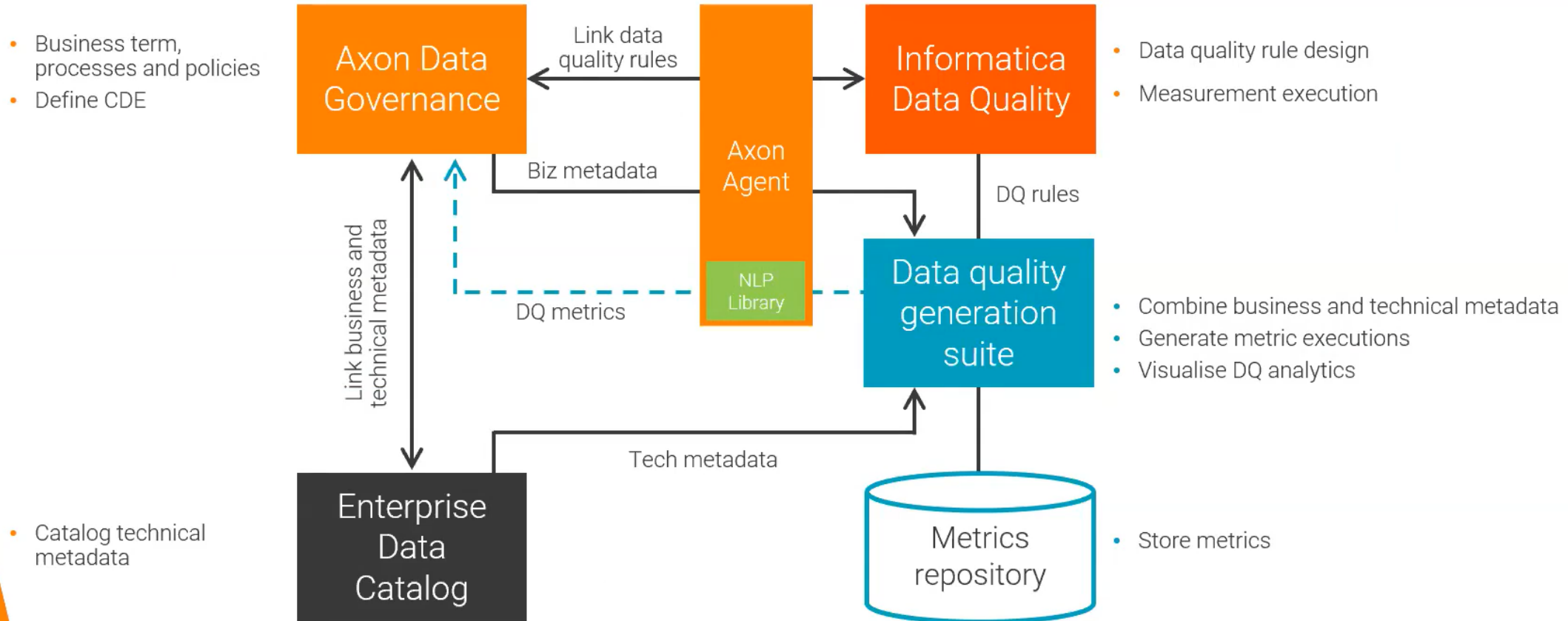


# Rule generation based on Business Context and Definition



# Automation Quality for Data Governance

## End to End Process Automation

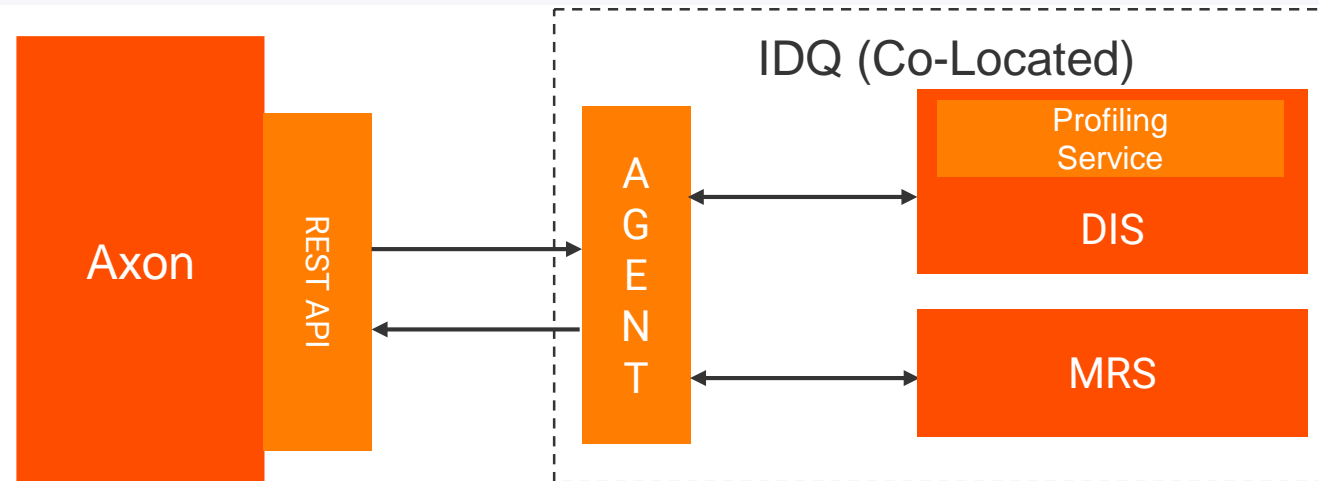


# Axon integration with IDQ

## Configuration

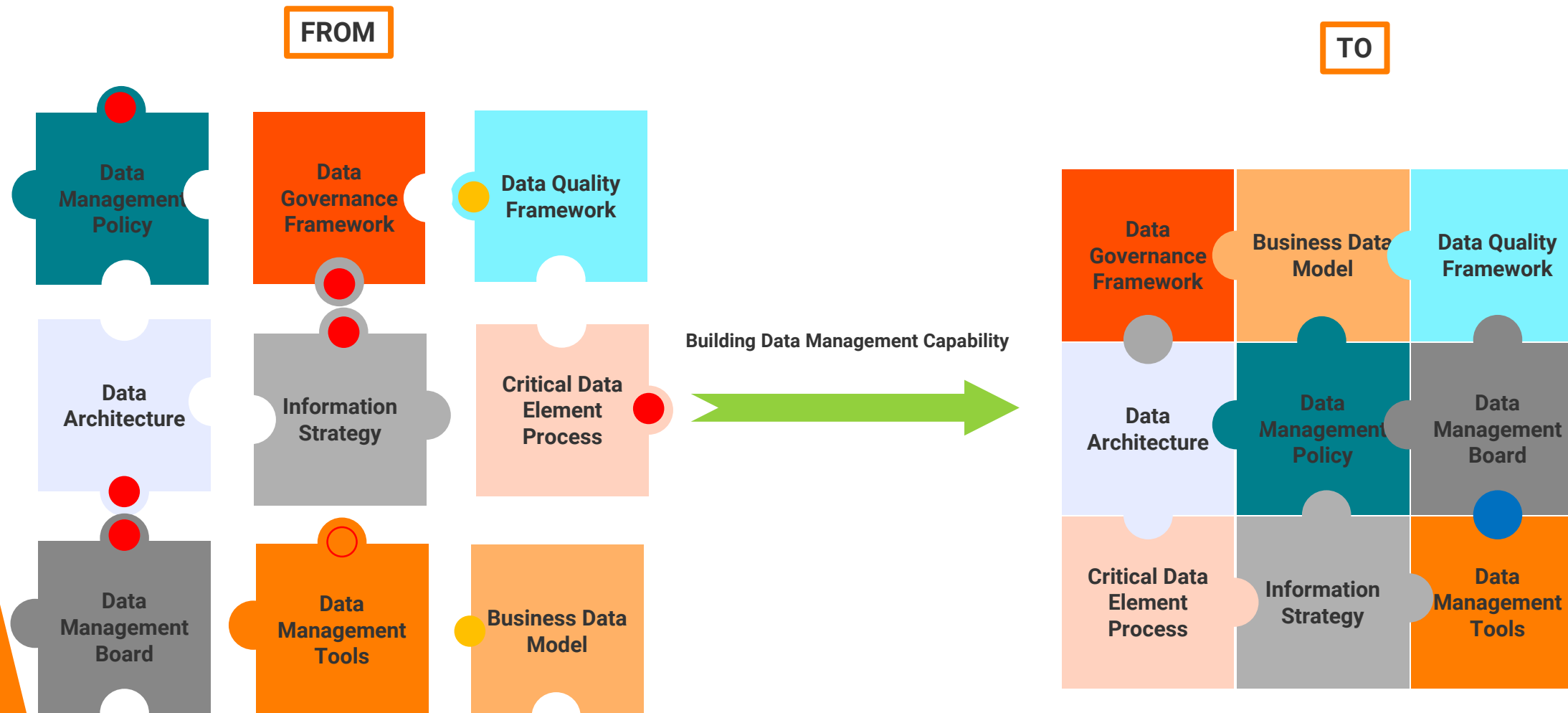
The screenshot shows the Informatica System Settings page for Axon Management. The left sidebar contains navigation links: Admin Dashboard, DG Operating Model, Meta-Model Administration, Operational Management, and Customize & Configure (selected). The main content area shows a 'Group' dropdown set to 'Data Quality' and an 'Edit' button. Below this is a table of system settings.

Display Name	Customizable	Value Type	Default Value	Value
Axon Agent Server Host Name ⓘ	✓	string		ascstaxon01.informatica.com
Axon Agent Server Port ⓘ	✓	string		10101



Demo

# Business Outcome



**ARE WE READY FOR THE ORGANIZATIONAL CHANGE?**



Thank You!

---

# References

- Axon Playbook  
[https://kb.informatica.com/whitepapers/4/PublishingImages/Informatica%20Axon%20-%20The%20Data%20Governance%20Playbook%20\(6.2%20Release\).pdf](https://kb.informatica.com/whitepapers/4/PublishingImages/Informatica%20Axon%20-%20The%20Data%20Governance%20Playbook%20(6.2%20Release).pdf)
- Using Axon with Enterprise Data Catalog  
<https://kb.informatica.com/h2l/HowTo%20Library/1/1150-UsingAxonDataGovernancewithEnterpriseDataCatalog-H2L.pdf>
- Axon Checklist to Automate Data Quality Rules  
<https://docs.informatica.com/data-quality-and-governance/axon-data-governance/h2l/checklist-to-automate-data-quality-rules-in-axon-data-governance/abstract.html>
- 7 Best Practices to Drive Data Catalog Adoption  
[https://www.informatica.com/lp/7-best-practices-to-drive-data-catalog-adoption\\_3712.html#fbid=PYYnGF22B9T](https://www.informatica.com/lp/7-best-practices-to-drive-data-catalog-adoption_3712.html#fbid=PYYnGF22B9T)
- Data Quality Accelerator  
<https://kb.informatica.com/faq/7/Pages/21/534447.aspx?myk=data%20quality%20accelerator%20guide>
- New Features in Axon 6.3  
<https://docs.informatica.com/data-quality-and-governance/axon-data-governance/6-3/release-guide/axon-6-3/new-features-in-version-6-3.html>