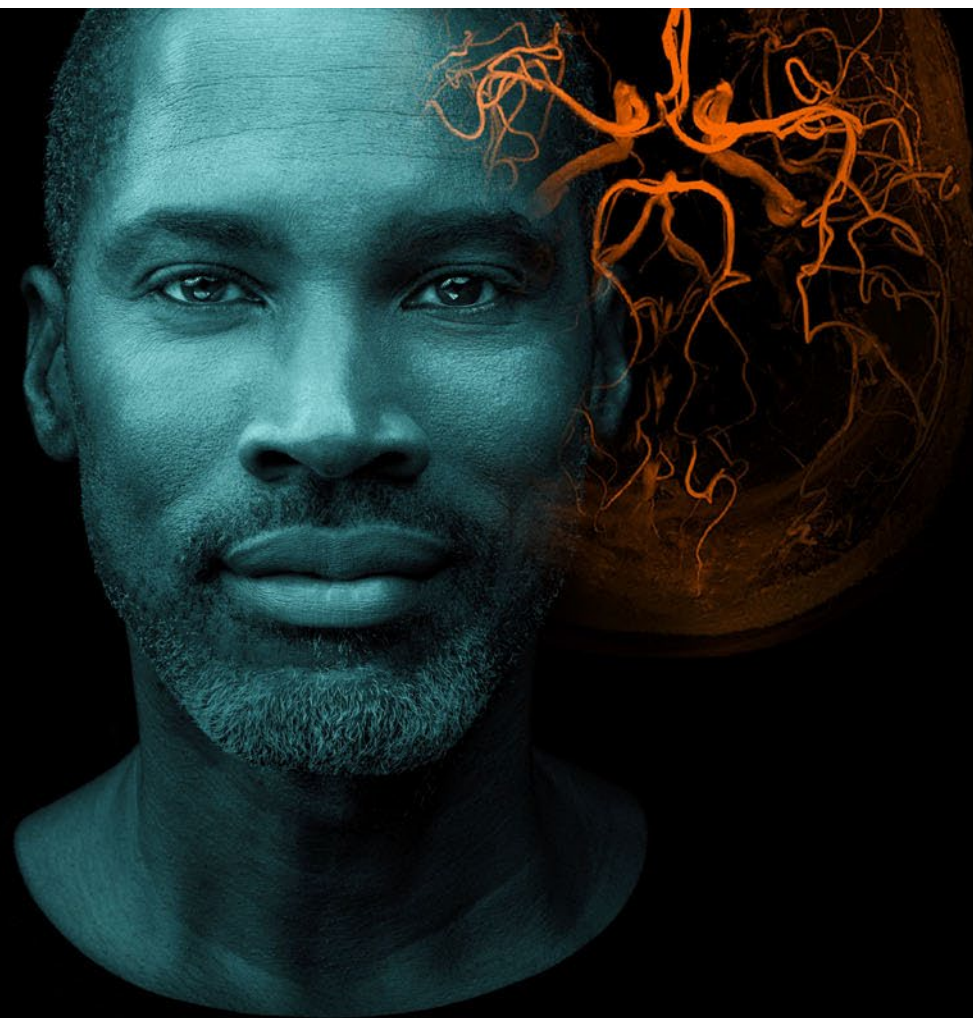
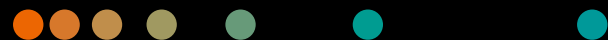


Can KNIME Accelerate your Data Governance Game?

KNIME Summit, 2024

Marcel Meyer – Siemens Healthineers



Many Data Governance initiatives face a common risk....



Source: Amazon.com

Spirits that I've cited
My commands ignore.

*Translated from the poem: "The Sorcerer's Apprentice" by Johann W. v. Goethe**

Here is your menu

1. Cheesy Crunch Galore
2. Savory Stack Surprise
3. Saucy Satisfaction Supreme
4. Midnight Masquerade Mystery
5. Enigmatic Euphoria Encounter

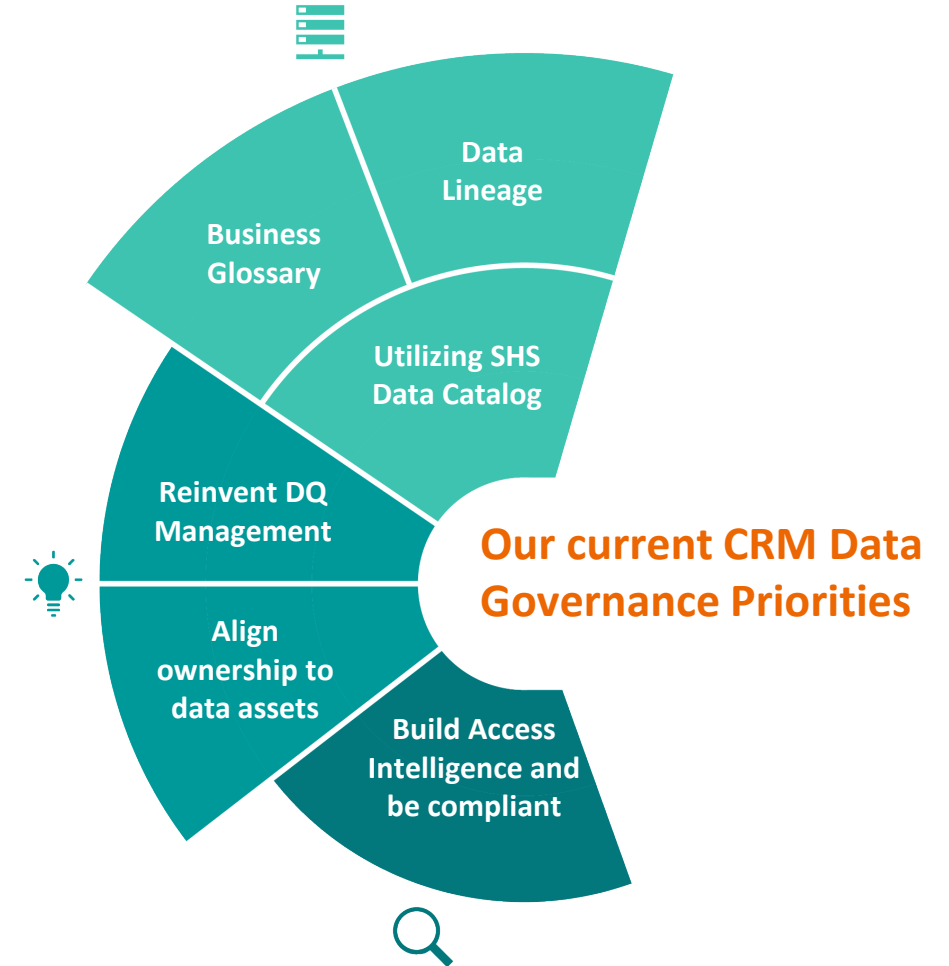
Our Burgers

- 1. Cheesy Crunch Galore**
Crispy Onion Rings, Melted cheddar cheese, Pickles, Coleslaw, BBQ sauce, Lettuce, Tomato
- 2. Savory Stack Surprise**
Smoked Gouda cheese, Sauteed spinach, Caramelized onions, Roasted garlic aioli, Arugula, Sun-dried tomatoes
- 3. Saucy Satisfaction Supreme**
Pepper jack cheese, Chipotle mayo, Grilled pineapple, Red cabbage slaw, Cilantro, Spicy jalapenos
- 4. Midnight Masquerade Mystery**
Juicy beef patty, Mysterious secret sauce, Smoked paprika dusting, Pickled red onions, Romaine lettuce, Sharp cheddar cheese, Crispy bacon
- 5. Enigmatic Euphoria Encounter**
Tender turkey patty, Enigmatic spice blend, Roasted garlic spread, Pickled jalapenos, Baby arugula, Pepper jack cheese, Avocado slices

Data Governance frameworks exist, but make it YOUR game

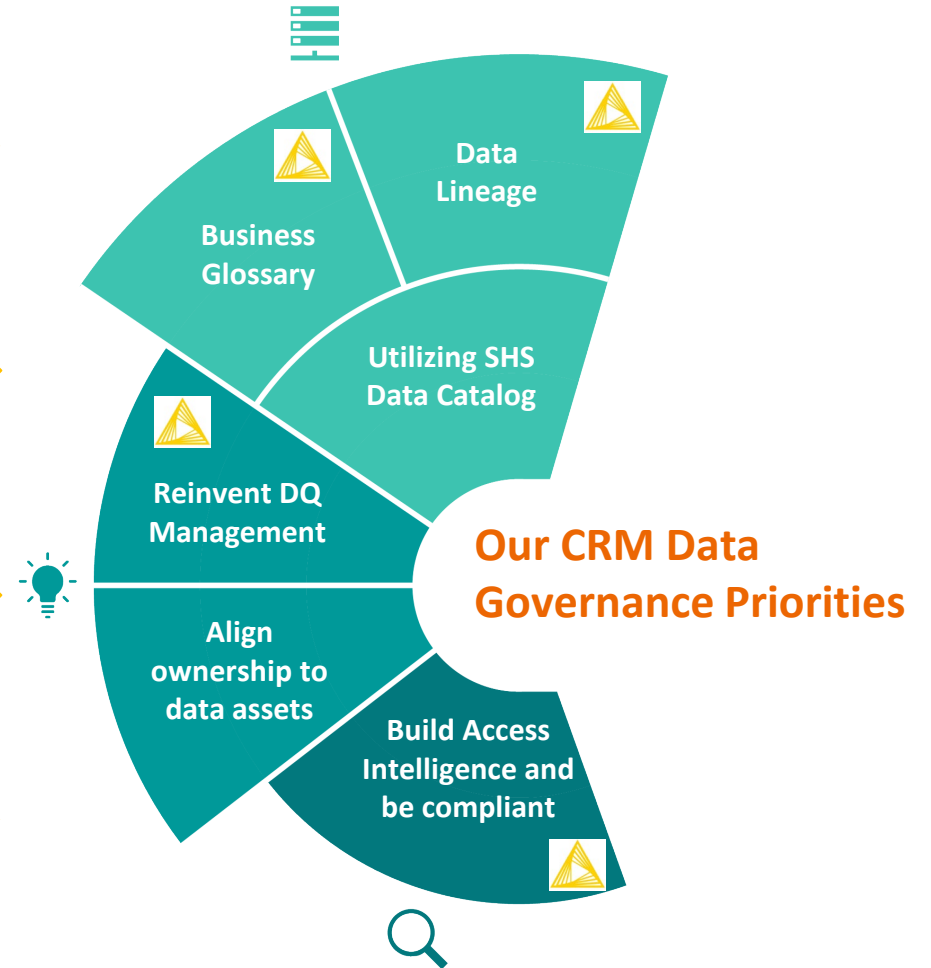
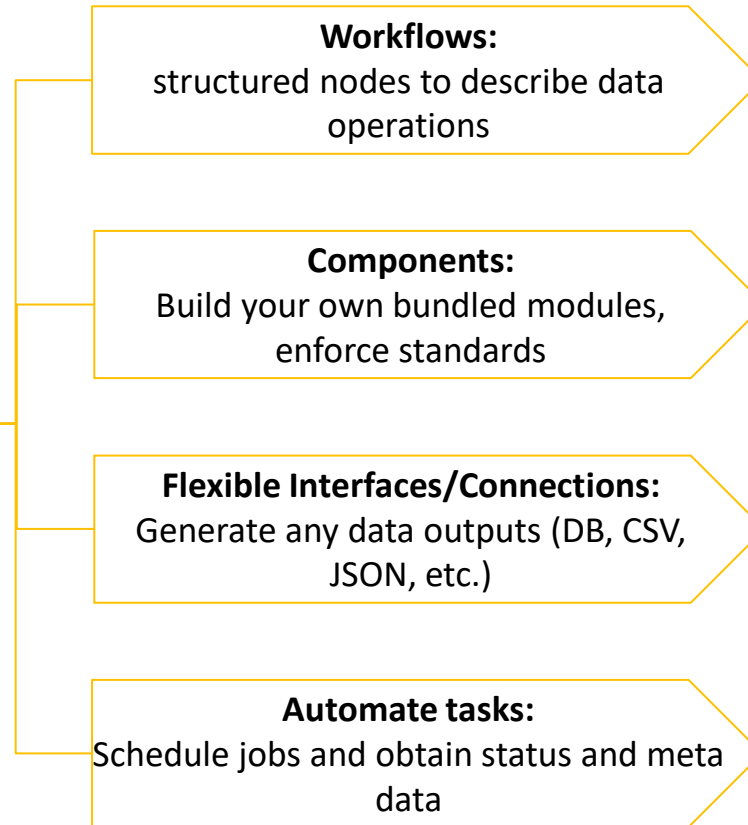
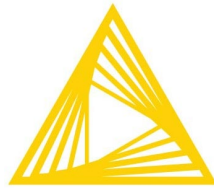


Copyright © DAMA International



Can KNIME Accelerate your Data Governance Game?

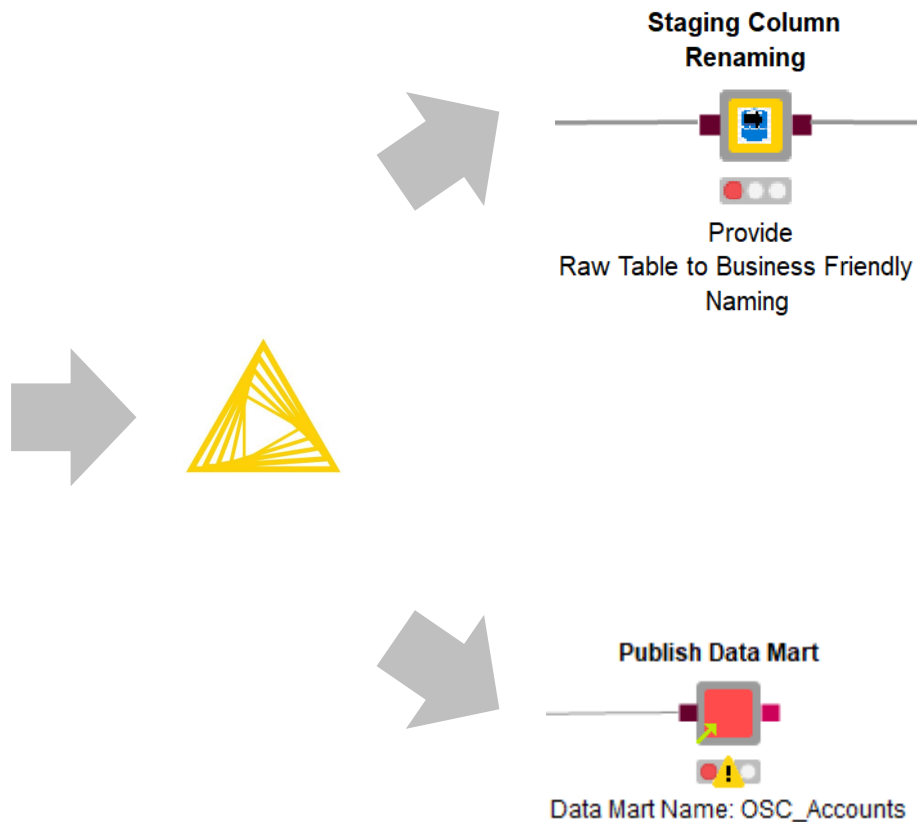
What does
KNIME offer?



Example Automated Meta Data for Glossary

Technical Column Name Source	Business Column Name	Description
COL_1	Column One	This represents value XYZ and is used for ABC

Centrally maintained repository of meta data



Columns: 301	DB...	D...	KNIME Type
ACCCLASSIFCTNINVITRO_CMEANING	VA...	V...	S String
ACCCLASSIFCTNINVIVO_CMEANING	VA...	V...	S String
ACCCLASSIFICATIONINVITRO_C	VA...	V...	S String
ACCLOCLATTRIBUTELOV1_CMEANING	VA...	V...	S String

Columns: 83	KNIME Type
ACCOUNT_ID	D...	...	D Number
Account Buying Group ID	V...	...	S String
Account Class In Vitro (DX)	V...	...	S String
Account Class In Vivo (EQ)	V...	...	S String
Account Country	V...	...	S String



Base Table with all Account information. Can be connected to any Table that provides the ACCOUNT_ID or a specific ACCOUNT_ID like ACCOUNT_ID (End User).

NAME ↑

OSC_Accounts hours ago

+

121 Columns

NAME ↑

ACCOUNT_ID

Identifies a individual Account Record from the CRM System

Example using KNIME workflow metadata to extract relevant data lineage information

Read Workflow & Job Metadata

Extract Input/Output & Data Operations

Format into Data Catalogue compatible format

```
JSON
{
  "version": "1.0.0",
  "summaryCreationDateTime": "2024-04-09 16:44:09 -0400",
  "environment": {},
  "workflow": {
    "nodes": [
      {
        "id": "240",
        "name": "Apply CDC Data Catalog",
        "type": "Manipulator",
        "state": "EXECUTED",
        "graphDepth": 3,
        "annotation": "Provide data lake table via table selector",
        "component": true
      }
    ]
  }
}
```

NodeFullName	ChildNodeFullName	Code
DB - Column Expression (Multi Column):0	DB - Column Expression (Multi Column):0	REGEXP_SUBSTR(\$CURRENT_COLUMN\$,PR-\\d*-\\d*)
DB - Column Expression (Multi Column):360	DB - Column Expression (Multi Column):360:330:334	CASE WHEN REGEXP_REPLACE(TRIM(\$CURRENT_CO...
DB - Column Expression (Multi Column):360	DB - Column Expression (Multi Column):360:336:0	to_decimal(floor(frull(\$CURRENT_COLUMN\$\$*Project...
DB - Column Expression (Multi Column):367	DB - Column Expression (Multi Column):367	CAST(\$CURRENT_COLUMN\$ AS DOUBLE)
DB - Rule Engine:360:359:341	DB - Rule Engine:360:359:341	'Project.First Rev. Recogn. before Signature Date (Flag...
DB - Rule Engine:360:359:340	DB - Rule Engine:360:359:340	('Project.Status' = 'Active' AND 'Project.Milestone Achie...
DB - Rule Engine:360:359:343	DB - Rule Engine:360:359:343	'Project.Status' = 'Active' AND 'Project.Milestone Achie...
DB - Rule Engine:360:359:344	DB - Rule Engine:360:359:344	'Project.Status' = 'Active' AND 'Project.Contract Signat...
DB - Rule Engine:360:359:352:348	DB - Rule Engine:360:359:352:348	'Project.Milestone Achieved'='NULL' => 0'Project.Miles...
DB - Rule Engine:360:335	DB - Rule Engine:360:335	'Project.Contract Value (GC)' > 5000000 => 'YTRUE =...
DB - Rule Engine:360:359:352:351	DB - Rule Engine:360:359:352:351	'Project.Contract Signature Date' < convert_timezone(...
DB - Rule Engine:360:359:352:349	DB - Rule Engine:360:359:352:349	'Project.Contract Signature Date' < convert_timezone(...
DB - Rule Engine:360:359:345	DB - Rule Engine:360:359:345	'Project.Status' = 'Active' AND 'Project.Milestone Achie...
DB - Rule Engine:360:337	DB - Rule Engine:360:337	'Project.Contract Value (LC)' > 5000000 => 'YTRUE =...
DB - Rule Engine:360:330:337	DB - Rule Engine:360:330:337	'Project.Other (LC)'>0 => 'Other'
DB - Rule Engine:360:330:336	DB - Rule Engine:360:330:336	'Project.Equipment VAR (LC)'>0 => 'HC VAR'



We could successfully onboard the data lineage into Informatica data catalogue making it available to our employees



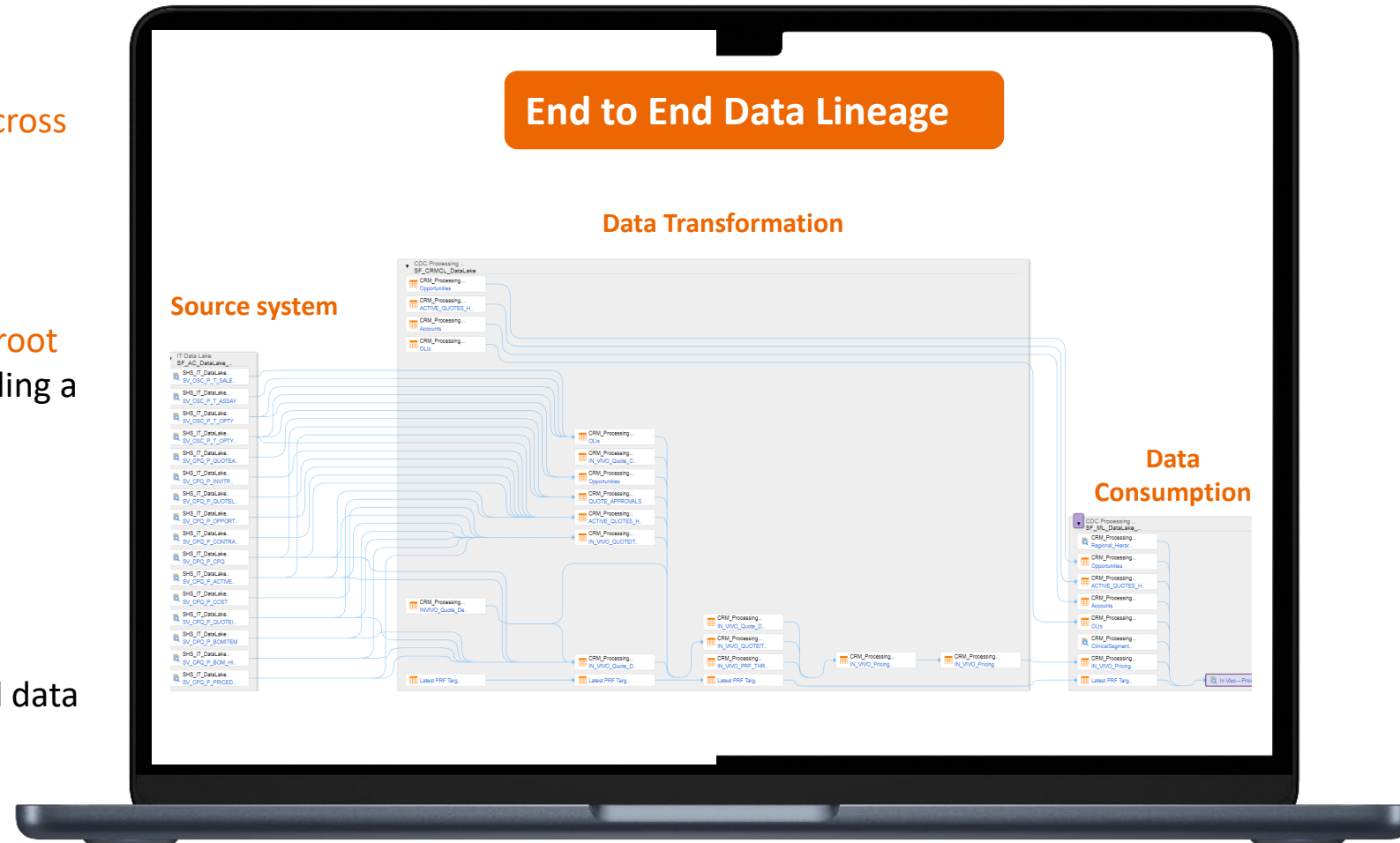
Data lineage provides a visual representation of SHS data's journey across multiple systems and transformations.



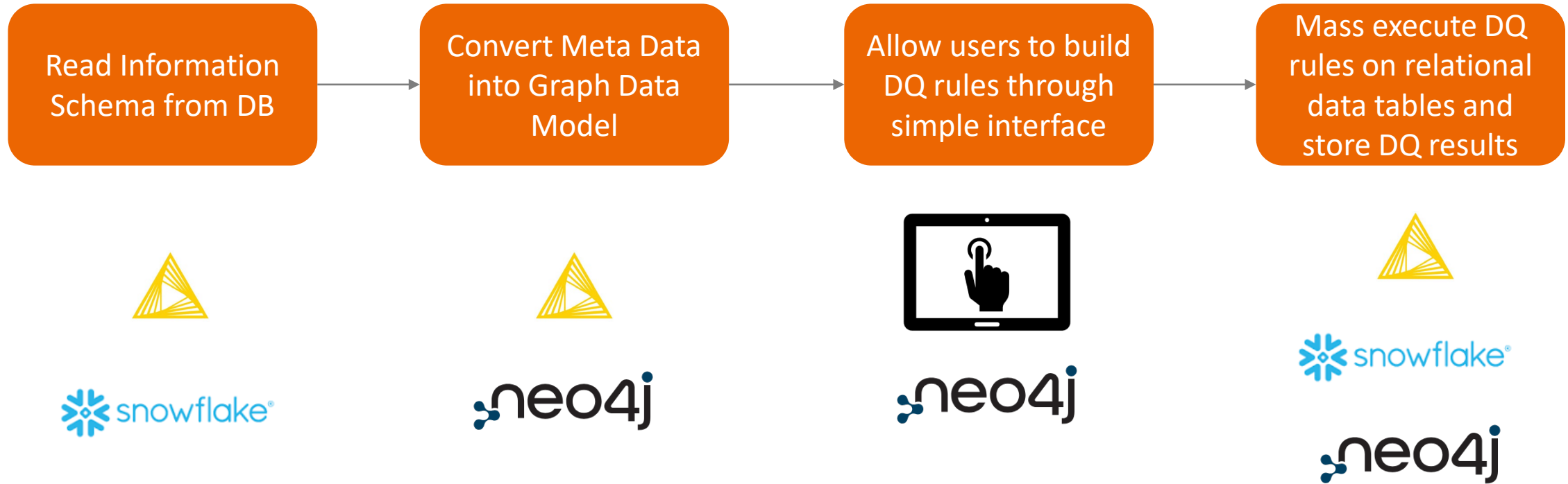
Data lineage facilitate impact analysis, root cause analysis and debugging by providing a clear view of the data's journey,



4000+ tables/views, 12000+ columns onboarded across multiple systems/applications downstream until data consumption.



Example DQM Automation - Overview



My key takeaways for you....

Data Governance
must add value

Meta Data shall be
brought as close
as possible to end
users

Smart data
workflow design
creates synergies
accelerating YOUR
DG game

Thank you! – Any questions?