

EDM Webinar

You need a data catalog. Do you know why?

A conversation with



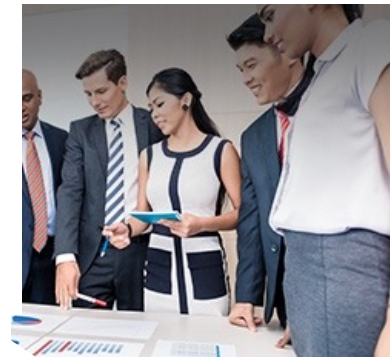
Matthew Vandever
VP, Strategic Services
Precisely



Christopher Reed
Manager, Sales Engineering
Precisely

precisely


EDM Council



Today's speakers

Moderator



Mike Meriton
Co-Founder & COO
EDM Council



Christopher Reed
Manager, Sales Engineering
Precisely



Matthew Vandevere
VP, Strategic Services
Precisely



The background is a solid deep purple. It features several 3D-rendered geometric shapes: a horizontal rectangle in the top left, a sphere to its right, a vertical rectangle in the top right, a sphere in the center, a sphere in the bottom left, a horizontal rectangle in the bottom left, and a vertical rectangle in the bottom right. All shapes have soft shadows, giving them a three-dimensional appearance.

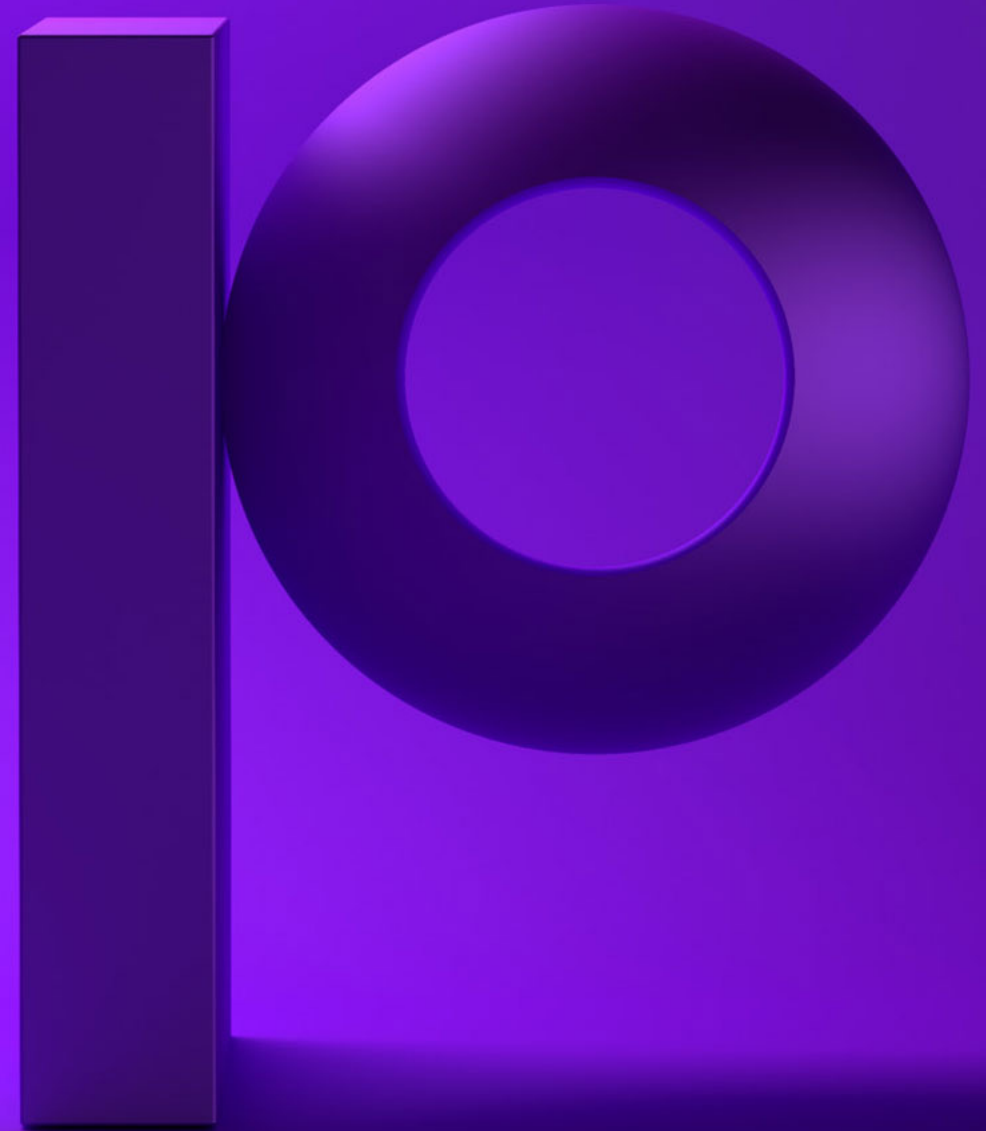
Poll #1

precisely



**You need a data catalog.
Do you know why?**

Christopher Reed| Manager, Sales Engineering
Matthew Vandever| VP, Strategic Services



Agenda

- What is a data catalog?
- The relationship between governance and the data catalog
- Data catalog governance use cases

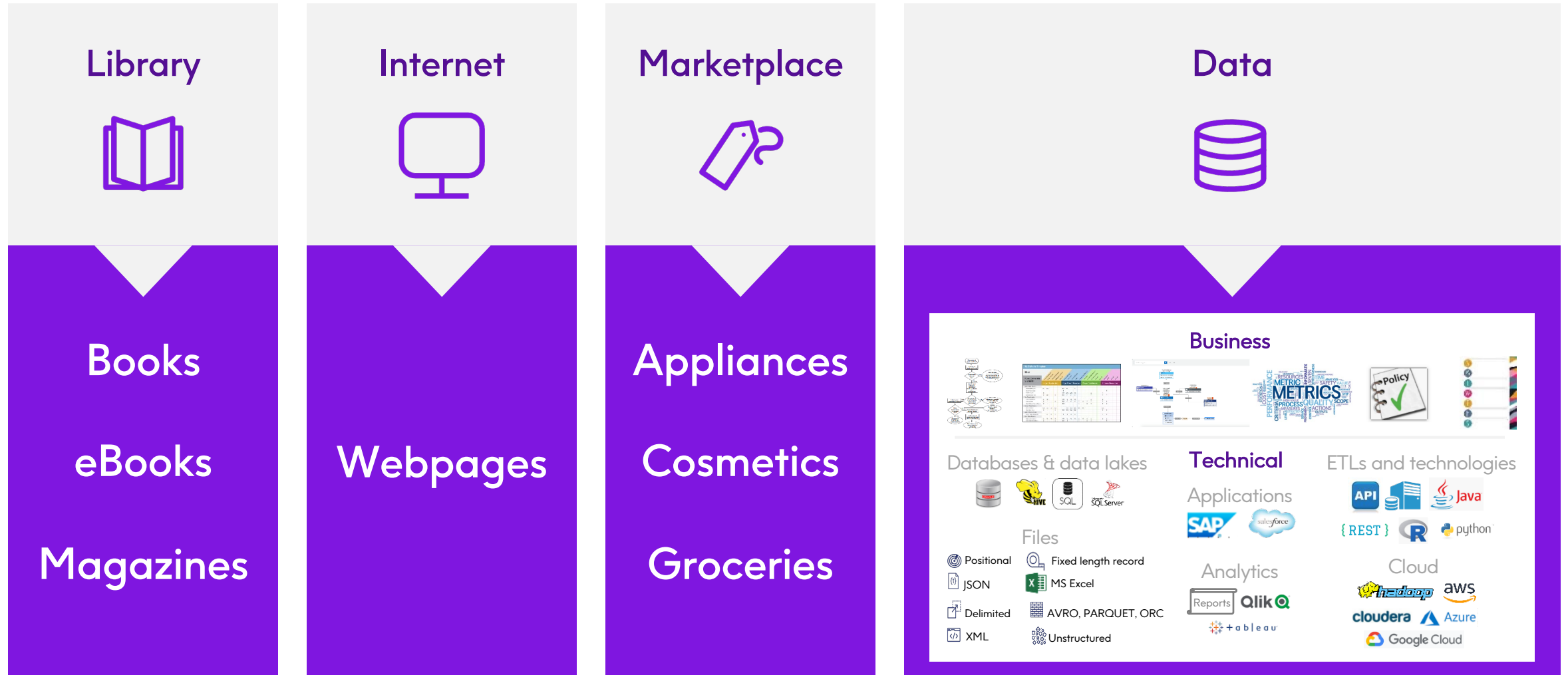
Data catalog drivers

Today's challenges

- We stood up a data lake, it's too much, now what?
- Profusion of valued analytics. Driving need for shareable data
- Regulatory compliance
- You've been told you need a data catalog
- My spreadsheet and email repositories are not keeping up



Data catalogs and data governance



Foundations of a catalog: consumer ready

	Why are you there?	What is it?	What is the outcome?	Who is responsible?	How is it maintained?	What are the rules?	How can I find something?	Where can I find it?
Library catalog	Research a topic or read book	Organized collection of books	Research a topic or read book	Librarian	Inventory control, shelving and shelf work	Books are available to the publish set by rules of behavior	Search by subject or peruse	Classification system. (subject-> shelf -> row)
Data catalog	Find a piece of data to solve a problem Research what my peers are using Monitor data and identify data impacts	Repository of data available to me with attribution, associations, relationships and lineage structured in a consistent metamodel	Request access and start working with it	Data steward The organization (everybody has a part to play)	Data stewardship (certification, scoring, quality, standards) Catalog itself (AI/ML learning, auto curation, workflow) Stakeholders / users / crowdsourced	Find a piece of data to solve a problem Research what my peers are using Monitor data and identify data impacts	Data is available to users in the organization according to data governance policy Data adheres to corporate policy (e.g. PII, standards, redundancy)	Asset path (system -> path -> element) by data domain

A data catalog enables business-ready data



Data that is ready to deliver outcomes

- Business objectives framework
- Enterprise governance & ownership
- Metrics & scoring



Data that is trusted

- Balancing, reconciliation & controls
- Stewardship & workflow
- Case management



Data that is easy to find and understand

- Data catalog & smart glossary
- Data lineage & impact analysis
- Data acquisition & analysis

What are the components of a “modern” data catalog?



Intuitive user
interface



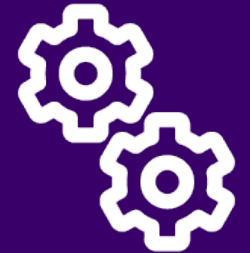
Has business
context



Contains extensive
information about
your data



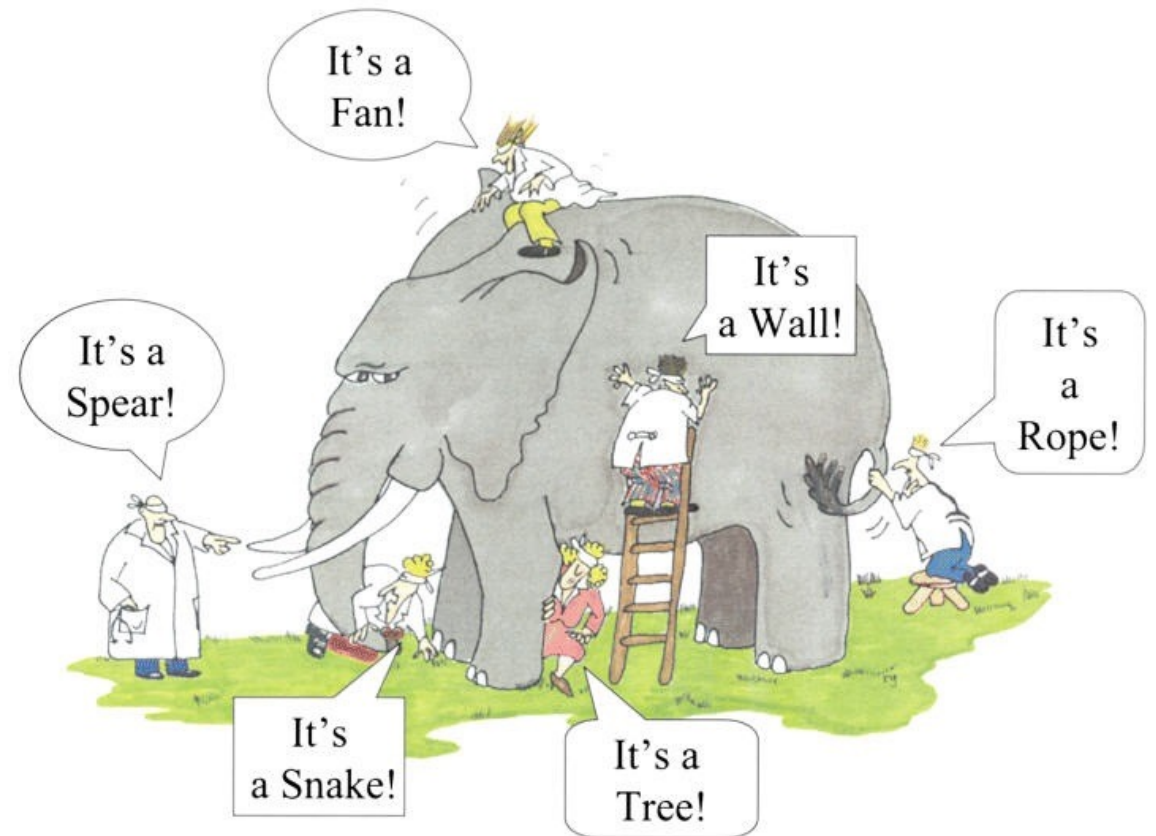
Promotes governance
and stewardship



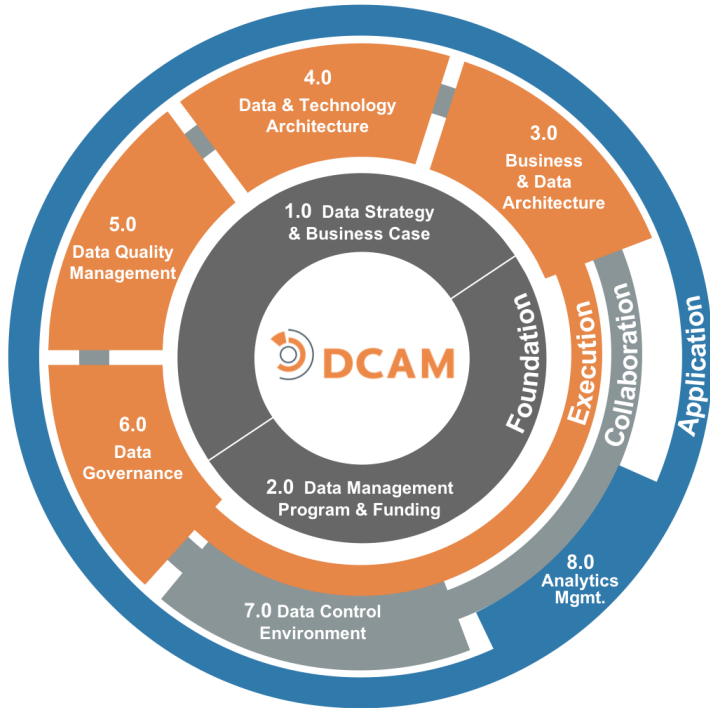
Turnkey automation,
administration,
and integration

What is governance?

Governance is often influenced by perspective



What is governance?

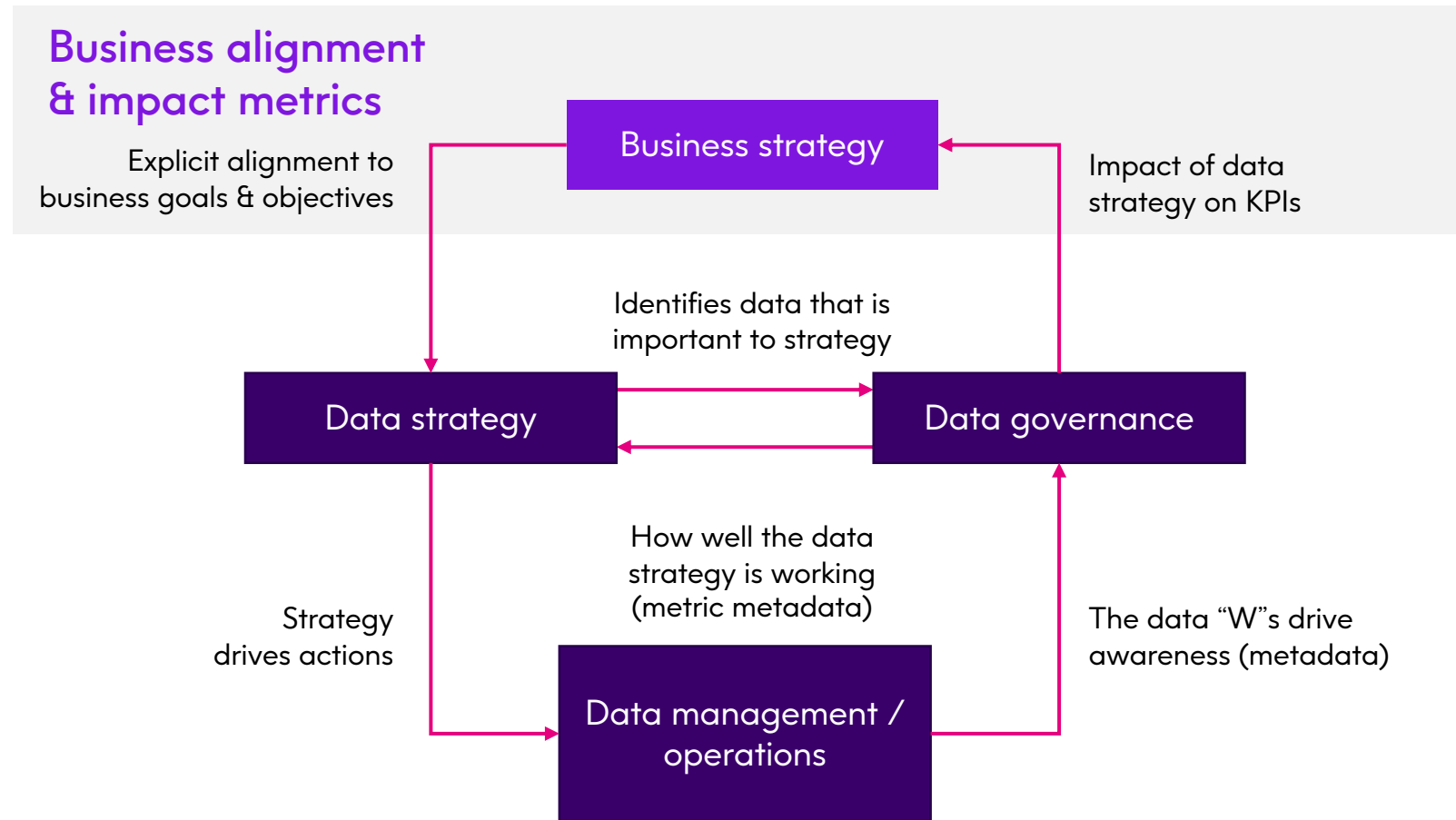


Governance = “The data catalog stores what good looks like and who is responsible”

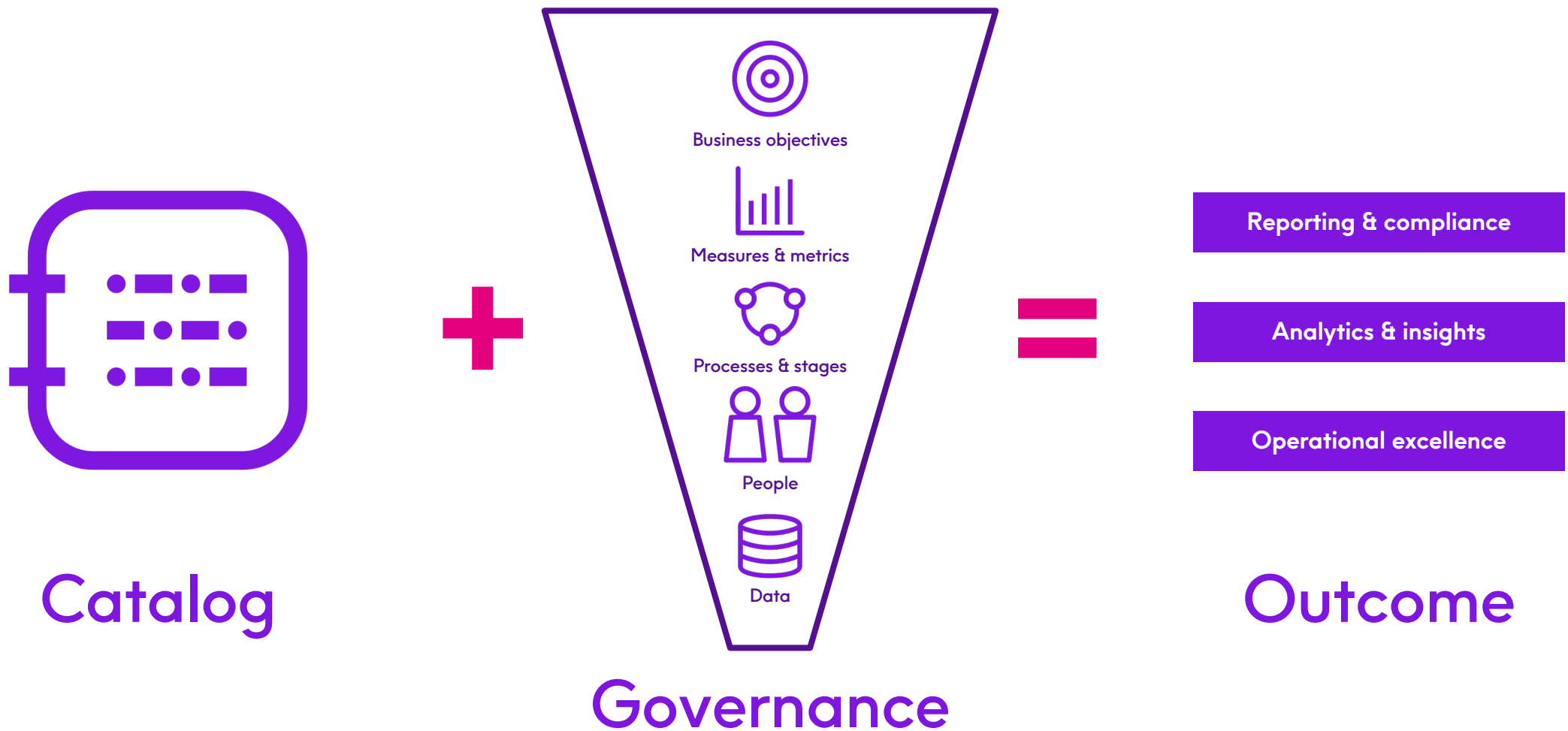
Governance activities

Data governance like other kinds of governance is focused on ensuring that activities are performed in accordance with strategy and accepted best practices in the organization

- In a well managed data environment data governance is well integrated with the business strategy
- Modern data governance does this proactively and passively
- Best in class solutions incorporate AI and ML to facilitate recommendation and automation engines

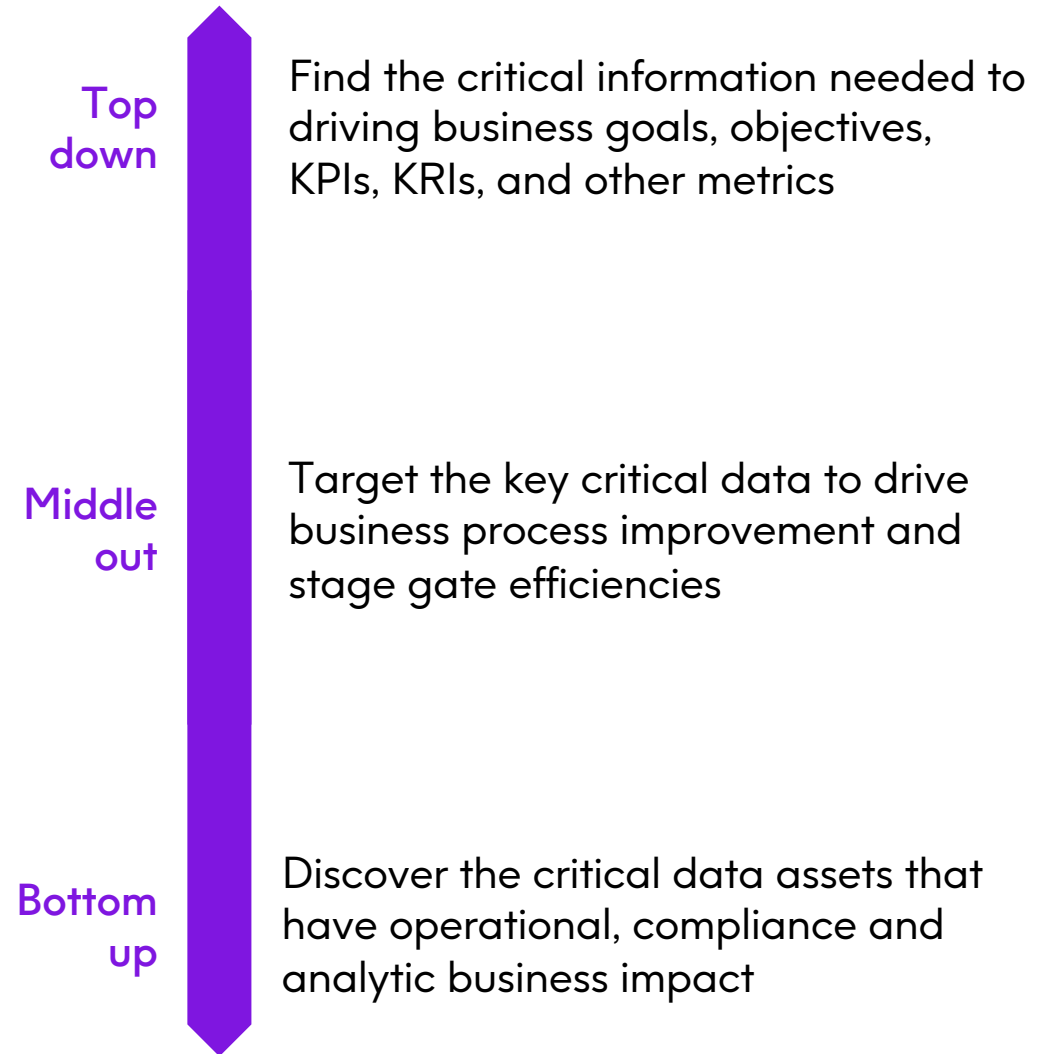


Data governance is the keystone for your data catalog

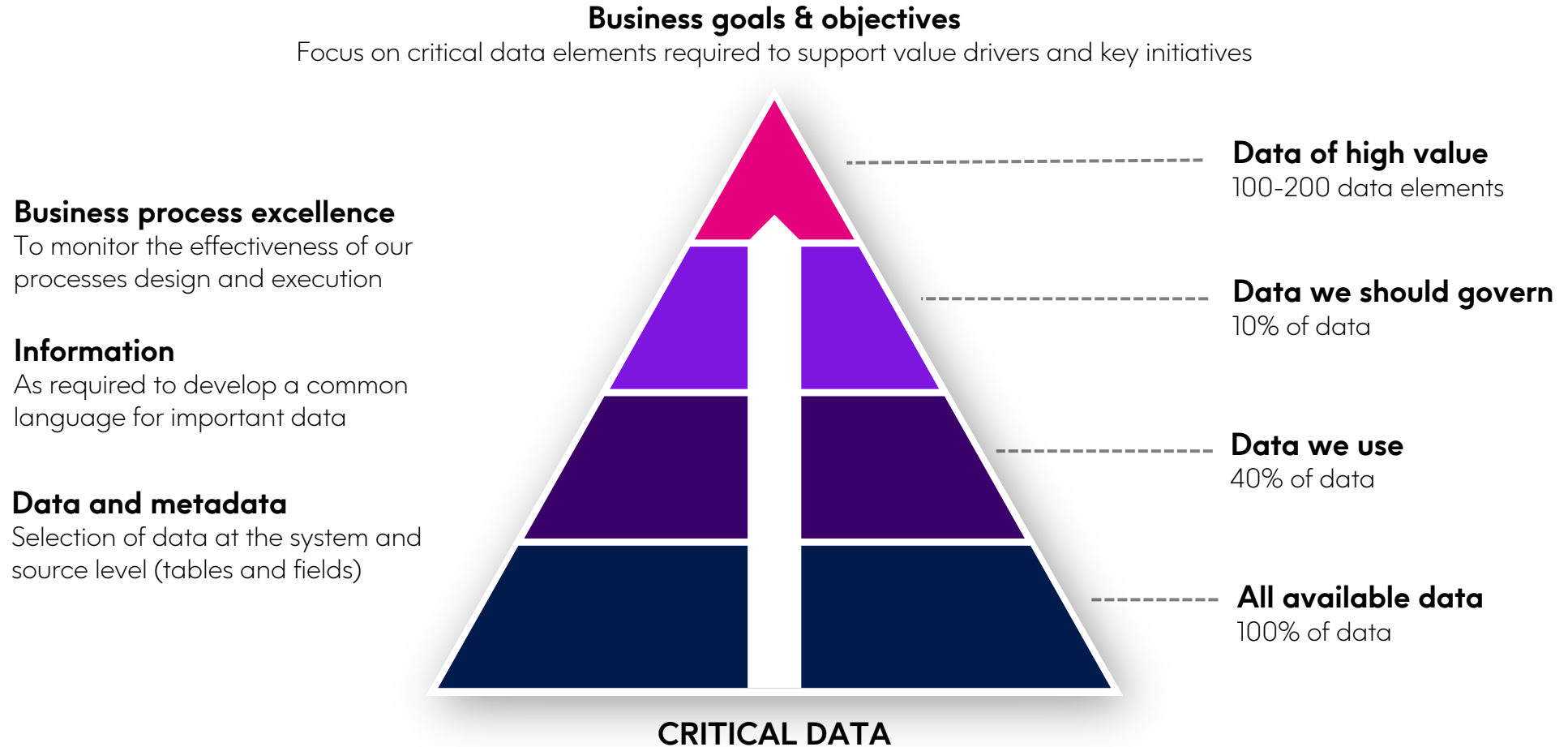


Governance follows a methodology

- Factors include: capabilities, culture, maturity
- All approaches assume automation and integration
- All should be supported by your data catalog
- Governance is integral to each one



Governance focuses on critical data



The evolution has been towards greater reliance on rich metadata as it captures the data's findability, usability and appropriateness within a particular context – implies well cataloged data

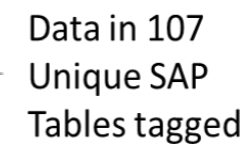
The background is a solid deep purple. It features several 3D-rendered geometric shapes: a horizontal rectangular prism at the top left, a sphere to its right, a small vertical rectangular prism in the center, a sphere to its right, a sphere at the bottom left, a square to its right, a sphere at the top right, and a tall vertical rectangular prism at the bottom right. All shapes have soft shadows and highlights, giving them a three-dimensional appearance.

Death by excel!

Getting off excel

```
SELECT Row as _,JAN ,FEB ,MAR, APR, MAY ,JUN ,JUL, AUG, SEP ,OCT ,NOV, DEC FROM (SELECT MAX(1) AS Sort ,MAX('Beginning of Month') AS Row
,Count(IIF(DATEDIFF('yyyy',RptYr,DateOpened)=-1,1)) AS JAN ,Count(Switch(Month(RptMth)<=1,NULL,(DATEDIFF('m',RptYr,DateOpened)<1 AND (DATEDIFF('m',RptYr,ClosedDate)<1
OR ClosedDate is Null)),1)) AS FEB ,Count(Switch(Month(RptMth)<=2,NULL,(DATEDIFF('m',RptYr,DateOpened)<2 AND (DATEDIFF('m',RptYr,ClosedDate)>=2 OR ClosedDate is Null)),1)) AS MAR ,Count(Switch(Month(RptMth)<=3,NULL,(DATEDIFF('m',RptYr,DateOpened)<3 AND (DATEDIFF('m',RptYr,ClosedDate)>=3 OR ClosedDate is Null)),1)) AS APR
,Count(Switch(Month(RptMth)<=4,NULL,(DATEDIFF('m',RptYr,DateOpened)<4 AND (DATEDIFF('m',RptYr,ClosedDate)>=4 OR ClosedDate is Null)),1)) AS MAY
,Count(Switch(Month(RptMth)<=5,NULL,(DATEDIFF('m',RptYr,DateOpened)<5 AND (DATEDIFF('m',RptYr,ClosedDate)>=5 OR ClosedDate is Null)),1)) AS JUN
,Count(Switch(Month(RptMth)<=6,NULL,(DATEDIFF('m',RptYr,DateOpened)<6 AND (DATEDIFF('m',RptYr,ClosedDate)>=6 OR ClosedDate is Null)),1)) AS JUL
,Count(Switch(Month(RptMth)<=7,NULL,(DATEDIFF('m',RptYr,DateOpened)<7 AND (DATEDIFF('m',RptYr,ClosedDate)>=7 OR ClosedDate is Null)),1)) AS AUG
,Count(Switch(Month(RptMth)<=8,NULL,(DATEDIFF('m',RptYr,DateOpened)<8 AND (DATEDIFF('m',RptYr,ClosedDate)>=8 OR ClosedDate is Null)),1)) AS SEP
,Count(Switch(Month(RptMth)<=9,NULL,(DATEDIFF('m',RptYr,DateOpened)<9 AND (DATEDIFF('m',RptYr,ClosedDate)>=9 OR ClosedDate is Null)),1)) AS OCT
,Count(Switch(Month(RptMth)<=10,NULL,(DATEDIFF('m',RptYr,DateOpened)<10 AND (DATEDIFF('m',RptYr,ClosedDate)>=10 OR ClosedDate is Null)),1)) AS NOV
,Count(Switch(Month(RptMth)<=11,NULL,(DATEDIFF('m',RptYr,DateOpened)<11 AND (DATEDIFF('m',RptYr,ClosedDate)>=11 OR ClosedDate is Null)),1)) AS DEC FROM (AccountMaster INNER JOIN UDF_AccountMaster ON AccountMaster.CustomerAccountNumber = UDF_AccountMaster.CustomerAccountNumber_Key) INNER JOIN ( SELECT *
,DATEADD('m',DATEDIFF('m',2,SystemDate)-1,2) as RptMth ,DATEADD('yyyy',DATEDIFF('yyyy',2,DATEADD('m',DATEDIFF('m',2,SystemDate)-1,2)),2) as RptYr FROM Control WHERE ClientName='MIL' ) AS Ctrl ON AccountMaster.ClientName = Ctrl.ClientName WHERE AccountMaster.ClientName = 'MIL' AND UDF_AccountMaster.ClientName = 'MIL' AND
(DATEDIFF('m',RptYr,ClosedDate)>-2 OR ClosedDate is Null) AND ManagerCode = '16S' GROUP BY AccountMaster.ClientName UNION SELECT MAX(2) AS Sort ,MAX('New Account') AS Row ,Count(IIF(Month(DateOpened)=1,1)) AS JAN ,Count(Switch(Month(RptMth)<=1,NULL,Month(DateOpened)=2,1)) AS FEB ,Count(Switch(Month(RptMth)
<=2,NULL,Month(DateOpened)=3,1)) AS MAR ,Count(Switch(Month(RptMth)<=3,NULL,Month(DateOpened)=4,1)) AS APR ,Count(Switch(Month(RptMth)
<=4,NULL,Month(DateOpened)=5,1)) AS MAY ,Count(Switch(Month(RptMth)<=5,NULL,Month(DateOpened)=6,1)) AS JUN ,Count(Switch(Month(RptMth)
<=6,NULL,Month(DateOpened)=7,1)) AS JUL ,Count(Switch(Month(RptMth)<=7,NULL,Month(DateOpened)=8,1)) AS AUG ,Count(Switch(Month(RptMth)
<=8,NULL,Month(DateOpened)=9,1)) AS SEP ,Count(Switch(Month(RptMth)<=9,NULL,Month(DateOpened)=10,1)) AS OCT ,Count(Switch(Month(RptMth)
<=10,NULL,Month(DateOpened)=11,1)) AS NOV ,Count(Switch(Month(RptMth)<=11,NULL,Month(DateOpened)=12,1)) AS DEC FROM (AccountMaster INNER JOIN UDF_AccountMaster ON AccountMaster.CustomerAccountNumber = UDF_AccountMaster.CustomerAccountNumber_Key) INNER JOIN ( SELECT * ,DATEADD('m',DATEDIFF('m',2,SystemDate)-1,2) as RptMth ,DATEADD('yyyy',DATEDIFF('yyyy',2,DATEADD('m',DATEDIFF('m',2,SystemDate)-1,2)),2) as RptYr FROM Control WHERE ClientName='MIL' ) AS Ctrl ON AccountMaster.ClientName = Ctrl.ClientName WHERE AccountMaster.ClientName = 'MIL' AND UDF_AccountMaster.ClientName = 'MIL' AND
(DATEDIFF('m',RptYr,ClosedDate)>-2 OR ClosedDate is Null) AND ManagerCode = '16S' GROUP BY AccountMaster.ClientName)
```

Answers need to be a few clicks away – not a few pivot tables away!



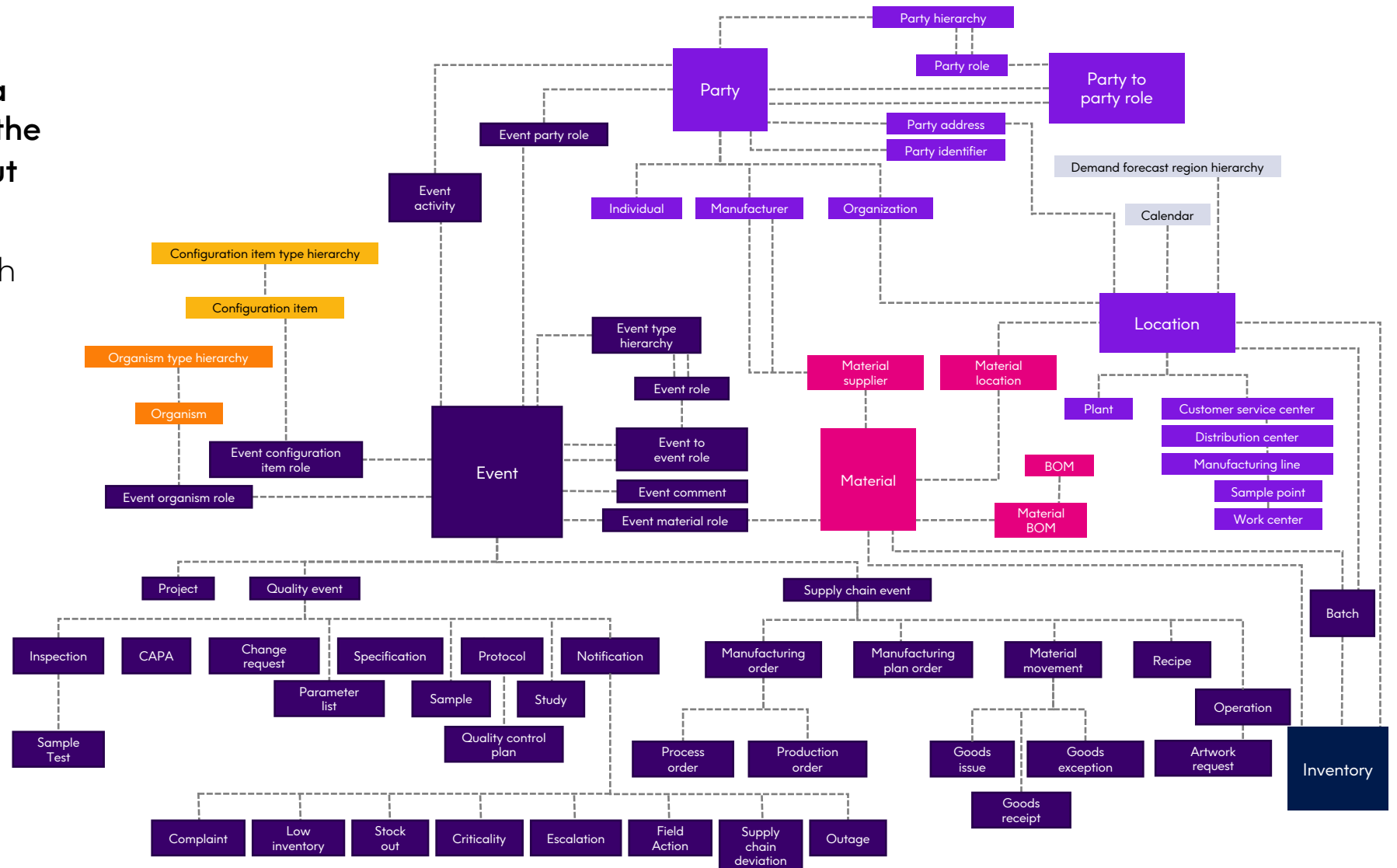
Catalog is a window to your data with many views!

A traditional enterprise data model can (maybe) handle the “single version of truth” – but not the many views of truth!

It is very hard to catalog both the data and the metadata around multiple use cases in Excel

How do I write my query to the EDW using the lens of:

- Risk
- ISO 27001
- HIPAA
- Privacy regulations (GDPR; CCPA)



An abstract composition on a dark purple background. It features several 3D-rendered geometric shapes: a horizontal rectangular prism at the top left, a sphere to its right, a vertical rectangular prism on the right side, a sphere above it, a sphere at the bottom left, and a square at the bottom center. The text 'Poll #2' is centered in the middle of the composition, and 'precisely' is at the bottom right.

Poll #2

precisely

Compliance

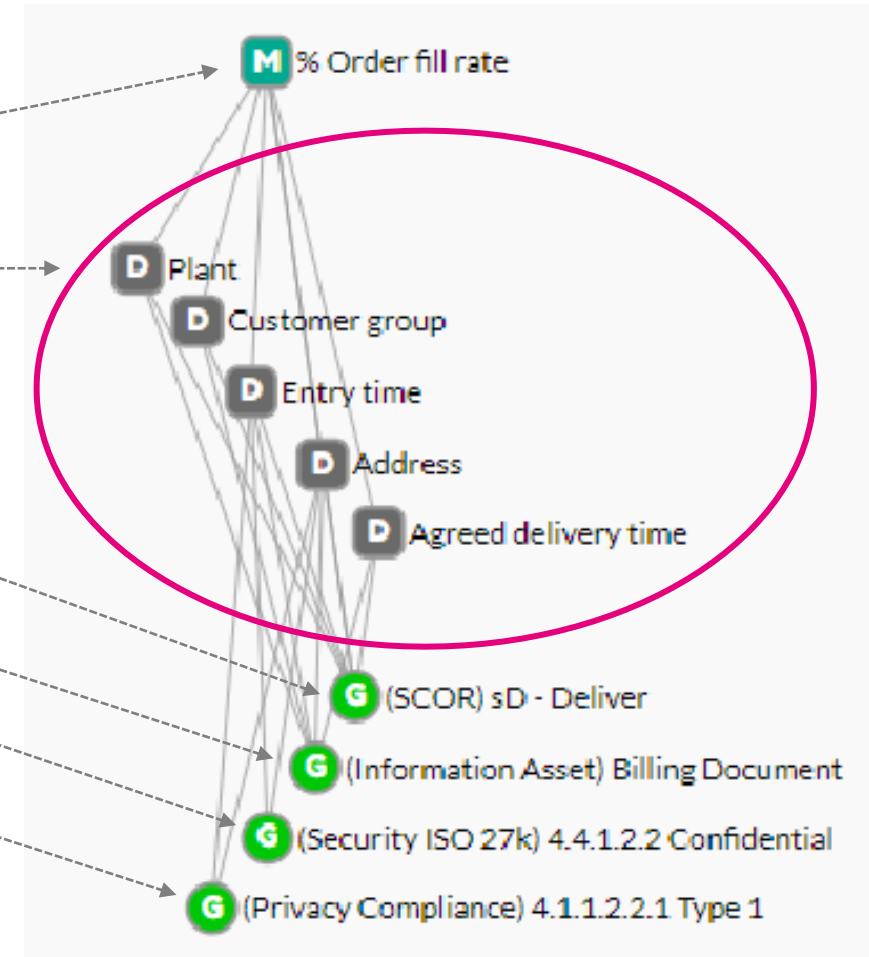
Catalogs for compliance...

If it is not cataloged – it is not governed!

Contained in catalog:

1. How and where used
2. Data is labelled to show where it is in the lifecycle
3. Data is linked to a data package
4. Data has security classification
5. Personal information “type” label flags this data as falling under privacy regulations

Do we know enough about the data to know it is being managed correctly?

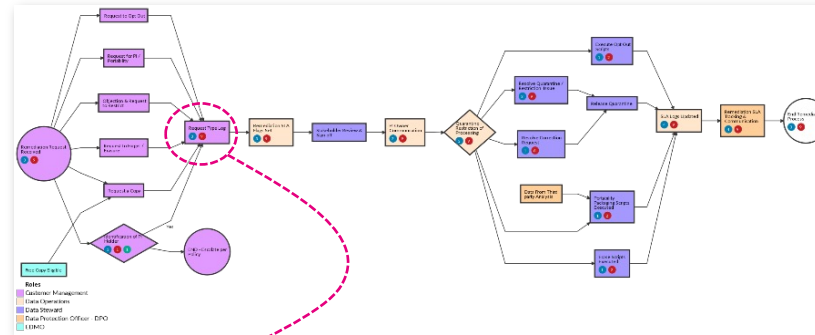


Audit control model has transparency and accountability

Accountability is defined for each control point...

Catalog contains the audit “view” of data:

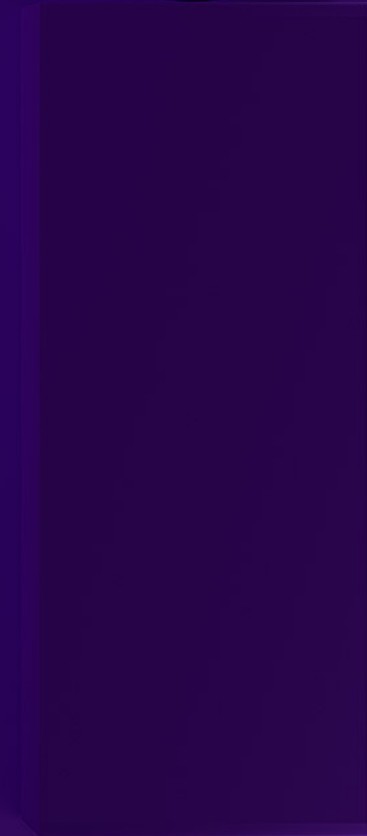
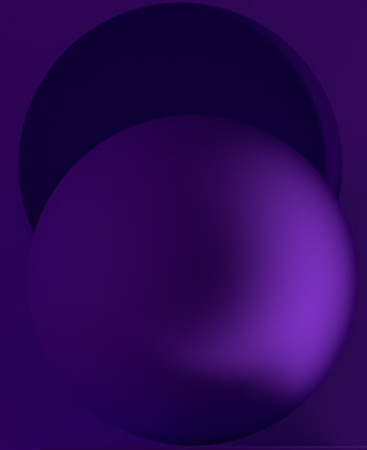
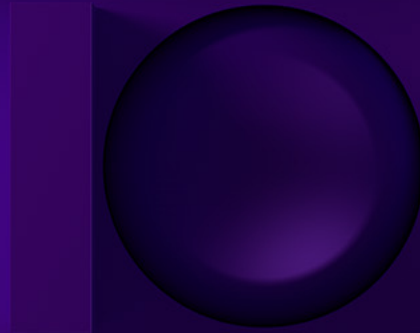
- A. Task owner
- B. Task detail
- C. The data required
- D. The standard that guides the execution
- E. The control rule that enforces the standard
- F. The metric that measures compliance



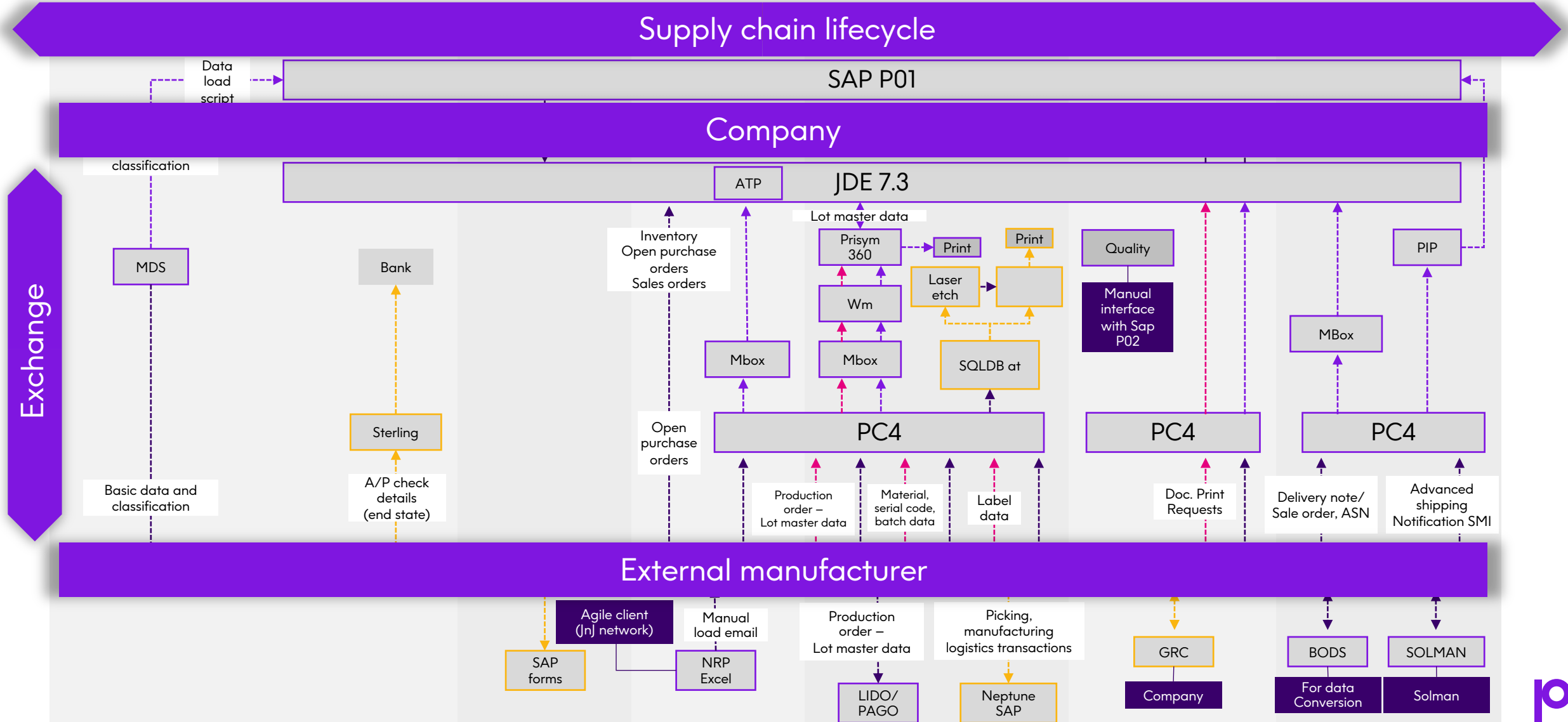
... which provides data level accountability

	A Task Owner	B Tasks	C Data	D Standard	E Control Rule	F Metric
70		Audit Log File Created	Z_DATUM GDPR DPO Contact Name (341) Z_DATUM GDPR DPO Contact email (342) Z_DATUM GDPR DPO Contact Phone (343) Z_DATUM GDPR DPO Published Flag 37-7 (344) Z_DATUM DPO Authority Org and Comm Review 38-1 to 6 (345) Z_DATUM GDPR DPO COI Date (346) Z_DATUM GDPR DPO COI Complete (347) Z_DATUM GDPR DPO COI Location (348) Z_DATUM GDPR Gov Framework CONOPS (349) Z_DATUM GDPR Gov Framework Data Check (351) Z_DATUM GDPR Gov Framework Standards Check (352) Z_DATUM GDPR Remediation Request Type (353)	Privacy Remediation Standard (13) Data Privacy Policy (3) DPO Policy (4) GDPR DPO Communication Plan (15)	GDPR DPO Designated 37-7 (14) GDPR DPO Published 37-7 (16) GDPR DPO Authority Org and Comm Review 38-1 to 6 (19) GDPR DPO COI Completeness 38-3; 38-6 (15) Gov Framework CONOPS Complete 25 - 1 to 3 (18) EDMO Gov Framework Complete (17) GDPR Remediation Request Type (12)	GDPR Remediation Logging Completeness (8)

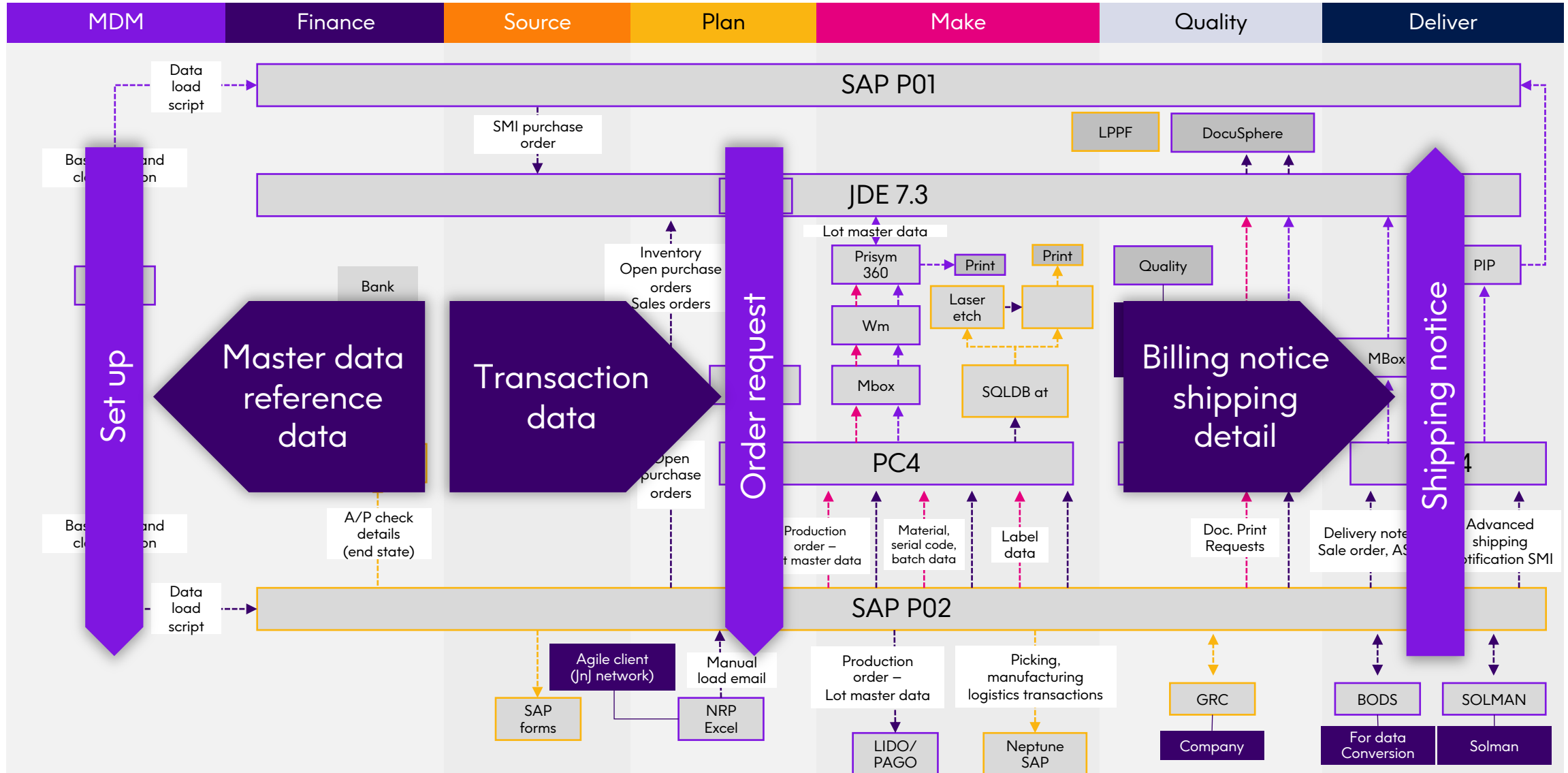
Optimizing operations



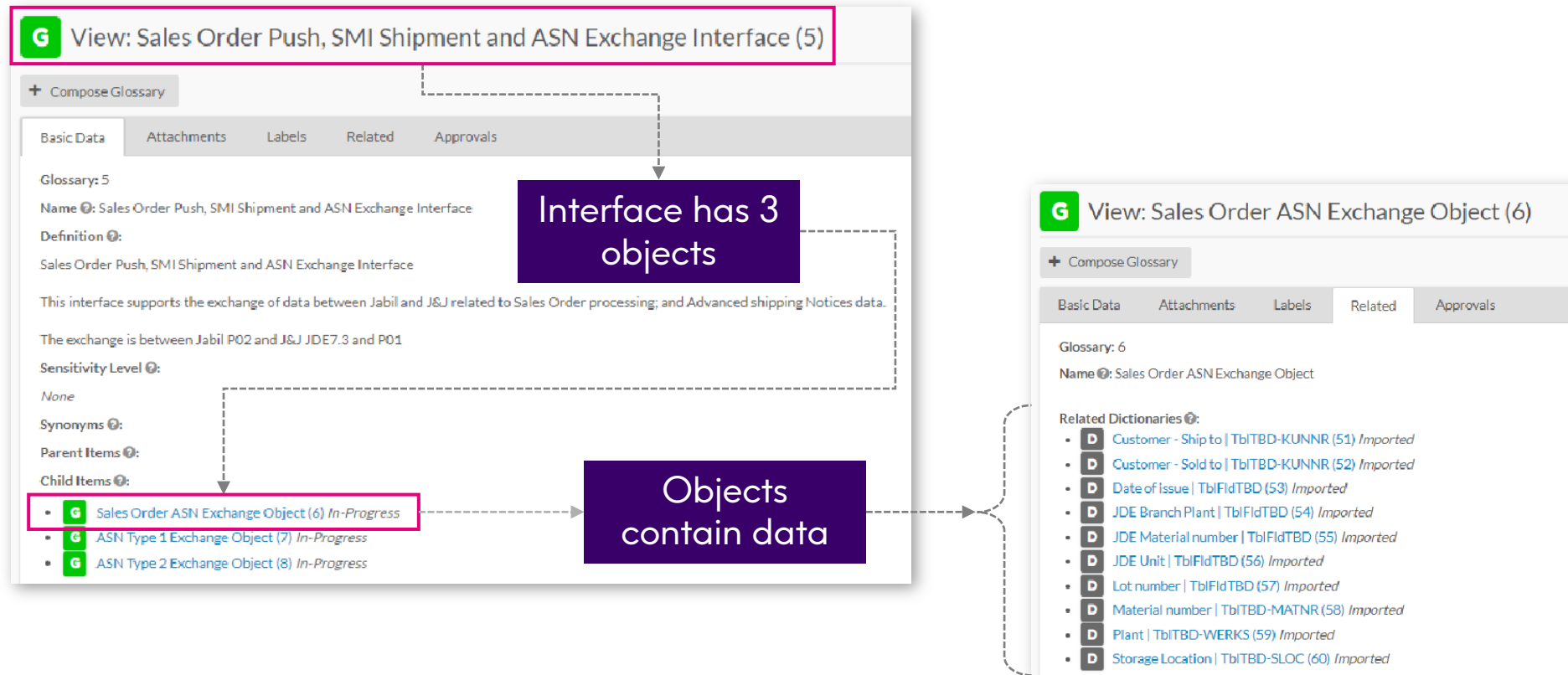
Data exchange with external manufacturing



Data exchange with external manufacturing



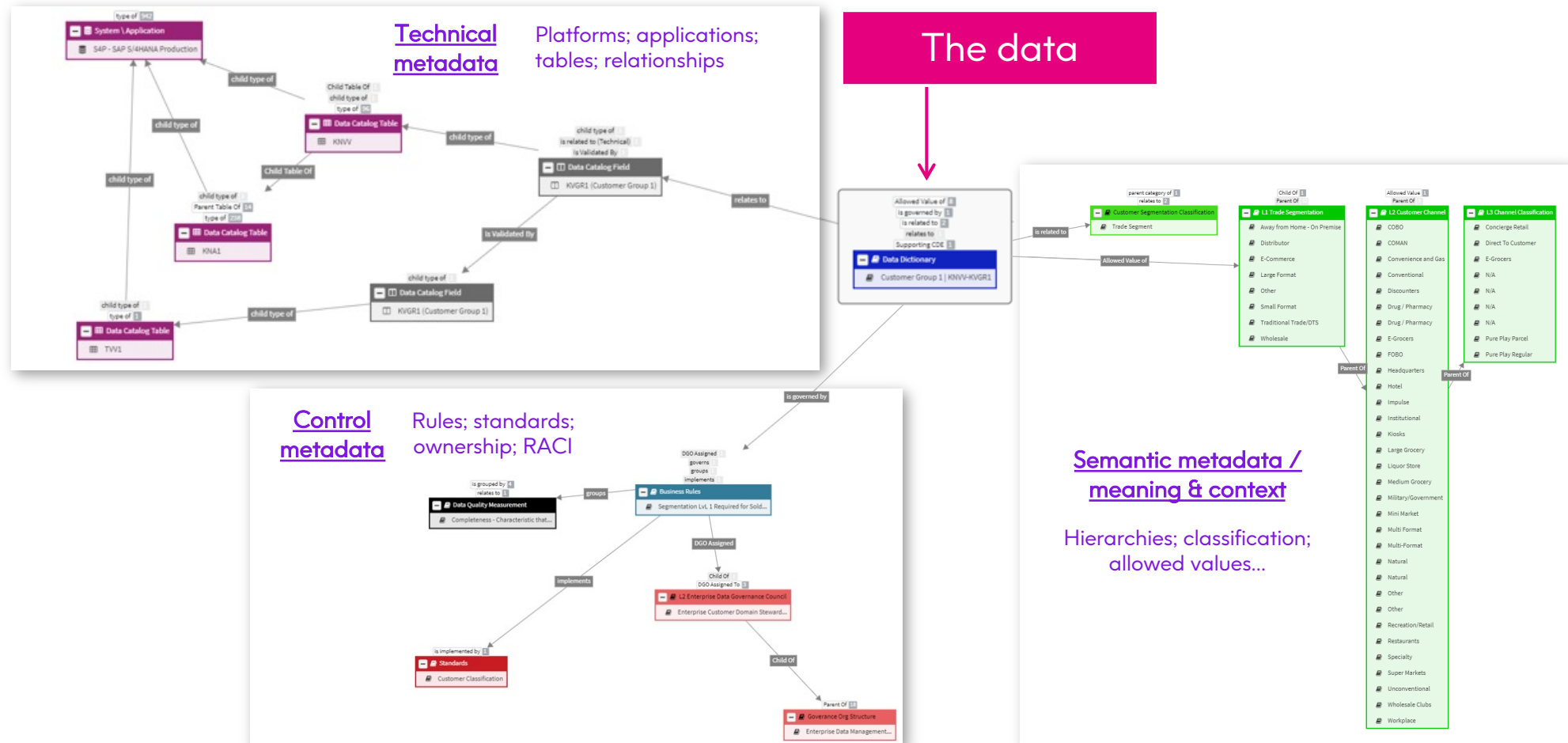
Cataloging captures issue PRIOR to production



Data quality rule: the shipping notice cannot contain any master data or reference data that was not contained in the set-up file

In closing... it is not just about the data!

How do you manage everything that surrounds the data? (metadata)



Questions?

Let's continue the conversation...



www.precisely.com

Demos
White Papers
Case Studies



Contact us

Set up a 30-minute personalized demo

Precisely.com/contact



precisely