

The Data Governance Platform Enablers

- Financial Services Journey Collecting, Cataloging and Disseminating Metadata

Member Since 1993
Former President

Founded in 1971



IRMAC
Information Resource Management
Association of Canada

Affiliated to DAMA International

Former Ethics Board Member



Associations *Experience*



Financial Services
Experience



Data Governance & Metadata Platforms *Experience*



Collibra

IBM InfoSphere Information Governance Catalog



Computer Associates

Ron Klein – I am a Data Operative

“In the discipline of Data Management, business acceptance is and has been a challenge for most, if not all practitioners. Conducting work nowadays with so many business strategies, compliance and risk agendas is even harder. Setting up Data Governance needs innovation and ample communication among stakeholders. One way out is being a Data Operative in its midst. While I am thinking on this approach and its several factors you should find it inspiring...”

What needs to be done!

Business metadata (e.g. definition, ownership, classification, policies/restrictions) about Data Domains and underlying critical data elements (CDEs) maintenance. Systems lineage including data flows for all CDEs

What needs to be done!

Business metadata (e.g. definition, ownership, classification, policies/restrictions) about Data Domains and underlying critical data elements (CDEs) maintenance. Systems lineage including data flows for all CDEs

Full end-to-end data lineage including attribute-level system lineage and business and technical transformation rules for Tier 1 CDEs

What needs to be done!

Business metadata (e.g. definition, ownership, classification, policies/restrictions) about Data Domains and underlying critical data elements (CDEs) maintenance. Systems lineage including data flows for all CDEs

Full end-to-end data lineage including attribute-level system lineage and business and technical transformation rules for Tier 1 CDEs

Data lineage and metadata information maintenance as changes occur and are reviewed at minimum annually on the Data Governance platform. Data Domain Stewards lead the maintenance of business metadata including domain-specific policy documents, while Data Custodians lead the maintenance of technical metadata

What needs to be done!

Business metadata (e.g. definition, ownership, classification, policies/restrictions) about Data Domains and underlying critical data elements (CDEs) maintenance. Systems lineage including data flows for all CDEs

Full end-to-end data lineage including attribute-level system lineage and business and technical transformation rules for Tier 1 CDEs

Data lineage and metadata information maintenance as changes occur and are reviewed at minimum annually on the Data Governance platform. Data Domain Stewards lead the maintenance of business metadata including domain-specific policy documents, while Data Custodians lead the maintenance of technical metadata

Critical data assigned classification which identifies confidentiality, protection, privacy, access and availability controls or restrictions

What needs to be done!

Business metadata (e.g. definition, ownership, classification, policies/restrictions) about Data Domains and underlying critical data elements (CDEs) maintenance. Systems lineage including data flows for all CDEs

Full end-to-end data lineage including attribute-level system lineage and business and technical transformation rules for Tier 1 CDEs

Data lineage and metadata information maintenance as changes occur and are reviewed at minimum annually on the Data Governance platform. Data Domain Stewards lead the maintenance of business metadata including domain-specific policy documents, while Data Custodians lead the maintenance of technical metadata

Critical data assigned classification which identifies confidentiality, protection, privacy, access and availability controls or restrictions

Data subject to ongoing assessment of key data risks categories. Preventative and detective controls to address and mitigate identified risks for CDEs

What needs to be done!

Business metadata (e.g. definition, ownership, classification, policies/restrictions) about Data Domains and underlying critical data elements (CDEs) maintenance. Systems lineage including data flows for all CDEs

Full end-to-end data lineage including attribute-level system lineage and business and technical transformation rules for Tier 1 CDEs

Data lineage and metadata information maintenance as changes occur and are reviewed at minimum annually on the Data Governance platform. Data Domain Stewards lead the maintenance of business metadata including domain-specific policy documents, while Data Custodians lead the maintenance of technical metadata

Critical data assigned classification which identifies confidentiality, protection, privacy, access and availability controls or restrictions

Data subject to ongoing assessment of key data risks categories. Preventative and detective controls to address and mitigate identified risks for CDEs

Data Quality defined and monitored for all CDEs to ensure data is fit for purpose. At minimum, data quality measures must cover the following dimensions: Accuracy, Completeness and Timeliness. Other dimensions (e.g. consistency, uniqueness, validity) can be measured where required

What needs to be done!

Business metadata (e.g. definition, ownership, classification, policies/restrictions) about Data Domains and underlying critical data elements (CDEs) maintenance. Systems lineage including data flows for all CDEs

Full end-to-end data lineage including attribute-level system lineage and business and technical transformation rules for Tier 1 CDEs

Data lineage and metadata information maintenance as changes occur and are reviewed at minimum annually on the Data Governance platform. Data Domain Stewards lead the maintenance of business metadata including domain-specific policy documents, while Data Custodians lead the maintenance of technical metadata

Critical data assigned classification which identifies confidentiality, protection, privacy, access and availability controls or restrictions

Data subject to ongoing assessment of key data risks categories. Preventative and detective controls to address and mitigate identified risks for CDEs

Data Quality defined and monitored for all CDEs to ensure data is fit for purpose. At minimum, data quality measures must cover the following dimensions: Accuracy, Completeness and Timeliness. Other dimensions (e.g. consistency, uniqueness, validity) can be measured where required

Issues related to critical data must be appropriately documented, tracked and remediated. Cross-functional and cross-domain data issues must be logged centrally

What needs to be done!

Business metadata (e.g. definition, ownership, classification, policies/restrictions) about Data Domains and underlying critical data elements (CDEs) maintenance. Systems lineage including data flows for all CDEs

Full end-to-end data lineage including attribute-level system lineage and business and technical transformation rules for Tier 1 CDEs

Data lineage and metadata information maintenance as changes occur and are reviewed at minimum annually on the Data Governance platform. Data Domain Stewards lead the maintenance of business metadata including domain-specific policy documents, while Data Custodians lead the maintenance of technical metadata

Critical data assigned classification which identifies confidentiality, protection, privacy, access and availability controls or restrictions

Data subject to ongoing assessment of key data risks categories. Preventative and detective controls to address and mitigate identified risks for CDEs

Data Quality defined and monitored for all CDEs to ensure data is fit for purpose. At minimum, data quality measures must cover the following dimensions: Accuracy, Completeness and Timeliness. Other dimensions (e.g. consistency, uniqueness, validity) can be measured where required

Issues related to critical data must be appropriately documented, tracked and remediated. Cross-functional and cross-domain data issues must be logged centrally

Maintain an inventory of approved Authoritative Sources (AS) and Authorized Provisioning Points (APP) which allow data to be shared efficiently with minimal impacts to source Systems of Record (SOR)

Data Vision

- Most Financial Services

Vision to become a data-driven organization that efficiently uses data to improve both customer experience and decision making



Value creation

Leverage data to build long term business advantage

- Optimize the cross-channel customer experience: anytime, anywhere
- Identify and act on cross-sell opportunities e.g., using AI
- Understand and predict customer needs, enhance relationships, and personalize experience
- Monetize data assets with third parties to create shared value



Governance

respond to regulators while building assets that generate insight

- Respond proactively to new data regulation
- Optimize the risks we take as an organization, balancing risk with potential value creation
- Reduce risk through effective risk management
- Ensure data is high quality
- Ensure safety, soundness, and privacy of customer data



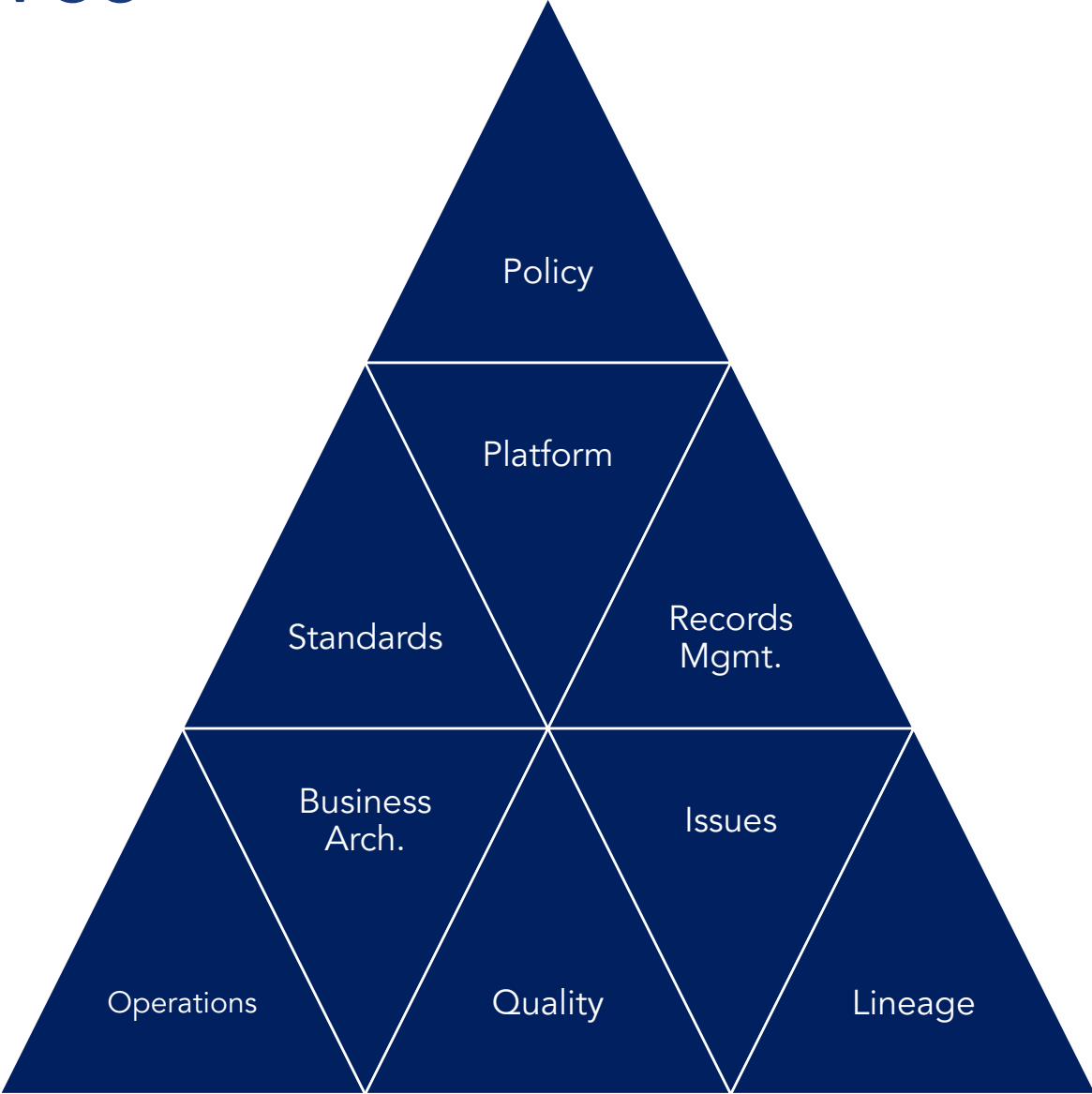
Operational excellence

manage data to reduce its Total Cost of Ownership

- Reduce data-related spend by leveraging a common and reusable data fabric - supporting a consistent approach to data across the enterprise
- Simplify and clarify data ownership and governance
- Ensure data is readily available (in the right form/granularity) to promote decision making

Data Initiatives

- *Everything Data*



Simplifying

Most Financial Organizations Goals



Easy to find, understand & trust data
drives reliable and accurate decision
making and insights

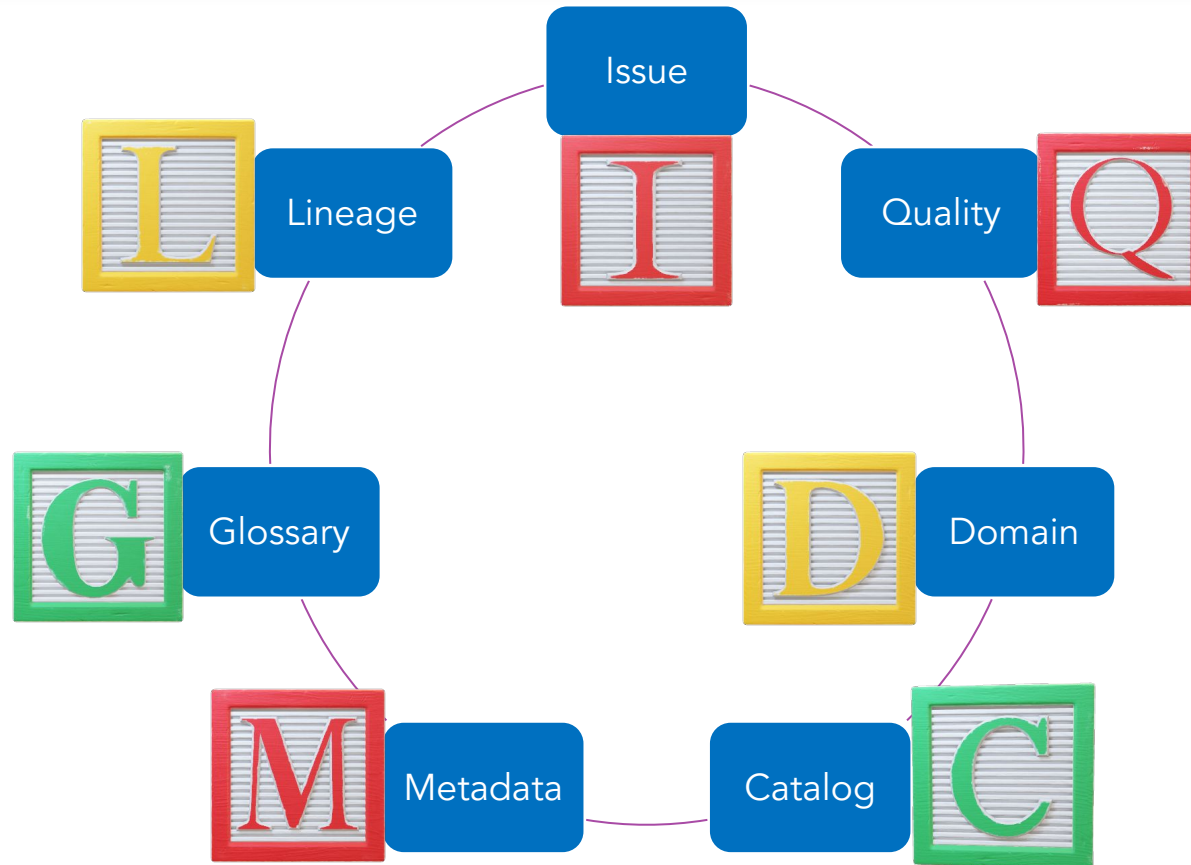


Easy to find, understand & trust data
reduces compliance and regulatory risk &
effort



Easy to find, understand & trust data
increases employee productivity

Data Governance – Alphabet Blocks



Alphabet Blocks – Definitions

- Issue – data issues – identification/ management / resolution
- Quality –Data Quality controls and process controls acting on data; where in the flow and importantly what are the results
- Domain – why data moves (business process, policies, risks)
- Lineage – how data moves across organization
- Catalog – link Glossary to Metadata - (book index: page (i.e. system/file/column) to find business terms)
- Metadata – physical characteristics of data in technical language (tables/columns/files/field)
- Glossary – common language/ taxonomy in business language so there is no ambiguity (business terms / CDE)

Data Governance – Triple Need



Business Glossary

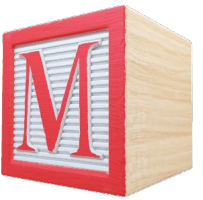
Find & Understand



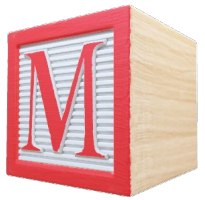
- Platform Features/Highlights
 - Governance/Ownership Structure
 - Workflows for creation, curation and modification
 - Staging/Reference Catalogs
 - CDE Tiering and Process context
 - Publishing/Integration to other tools
 - Reporting – Progress & Quality Exceptions
- Business/Process
 - Glossary guidelines and processes
 - Curation Team
 - Operational support/training
- Ball parking numbers
 - 4000 + reference terms
 - + 1/2 curated terms
 - + 10% CDE's

Metadata

Find & Understand

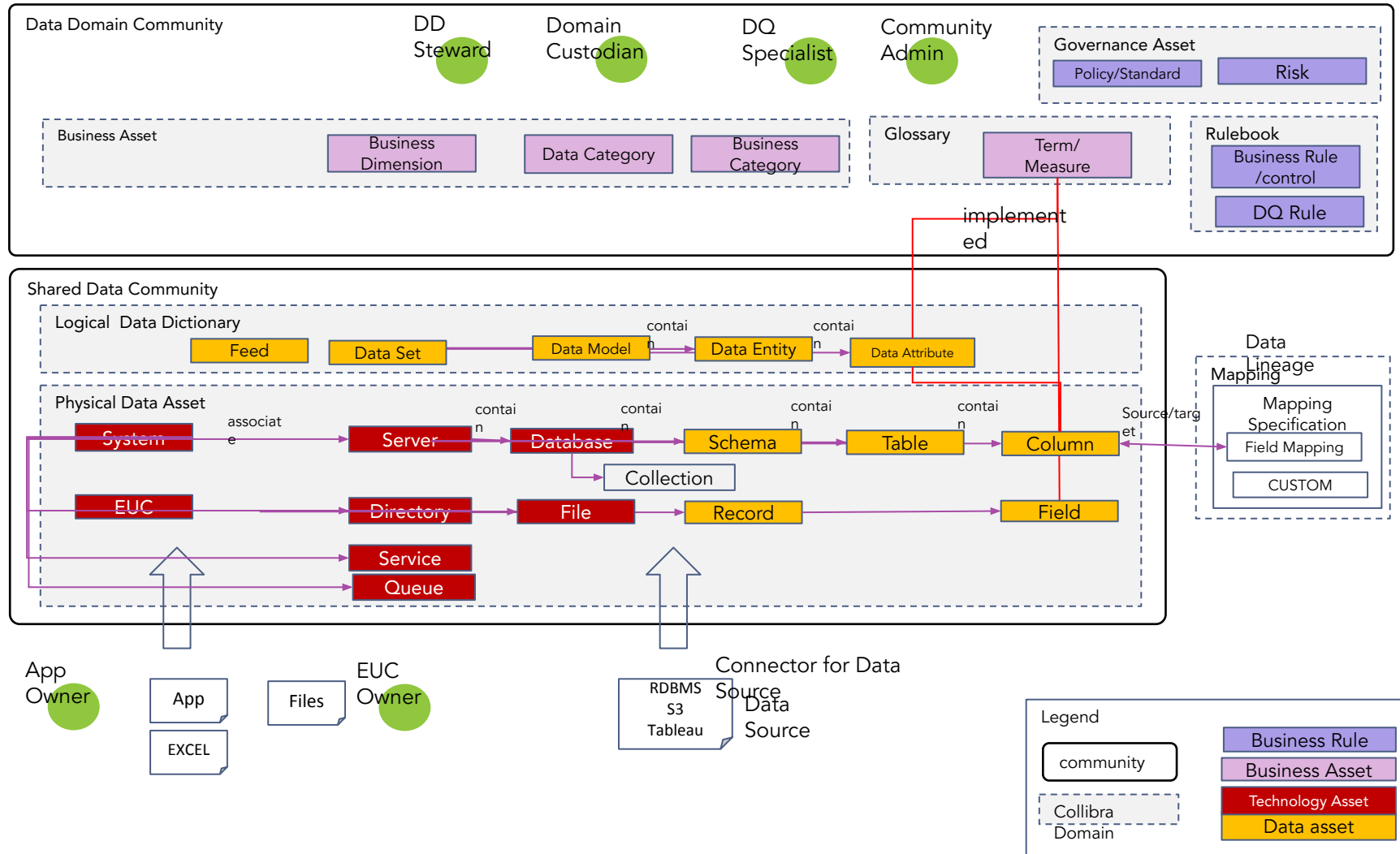


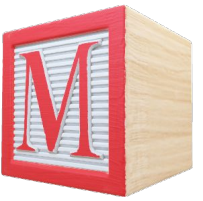
- Platform Features/Highlights
- Governance/Ownership Structure
- Naming Standards / Hierarchies
- Manual Load Templates
- Automated Metadata Framework – RDBMS Loads; Lake Loads
- Future: Technologies ++ (Tableau, S3 ...)
- Business/Process
- Metadata Standards
- Operational guidelines and processes
- Support/training



Metadata

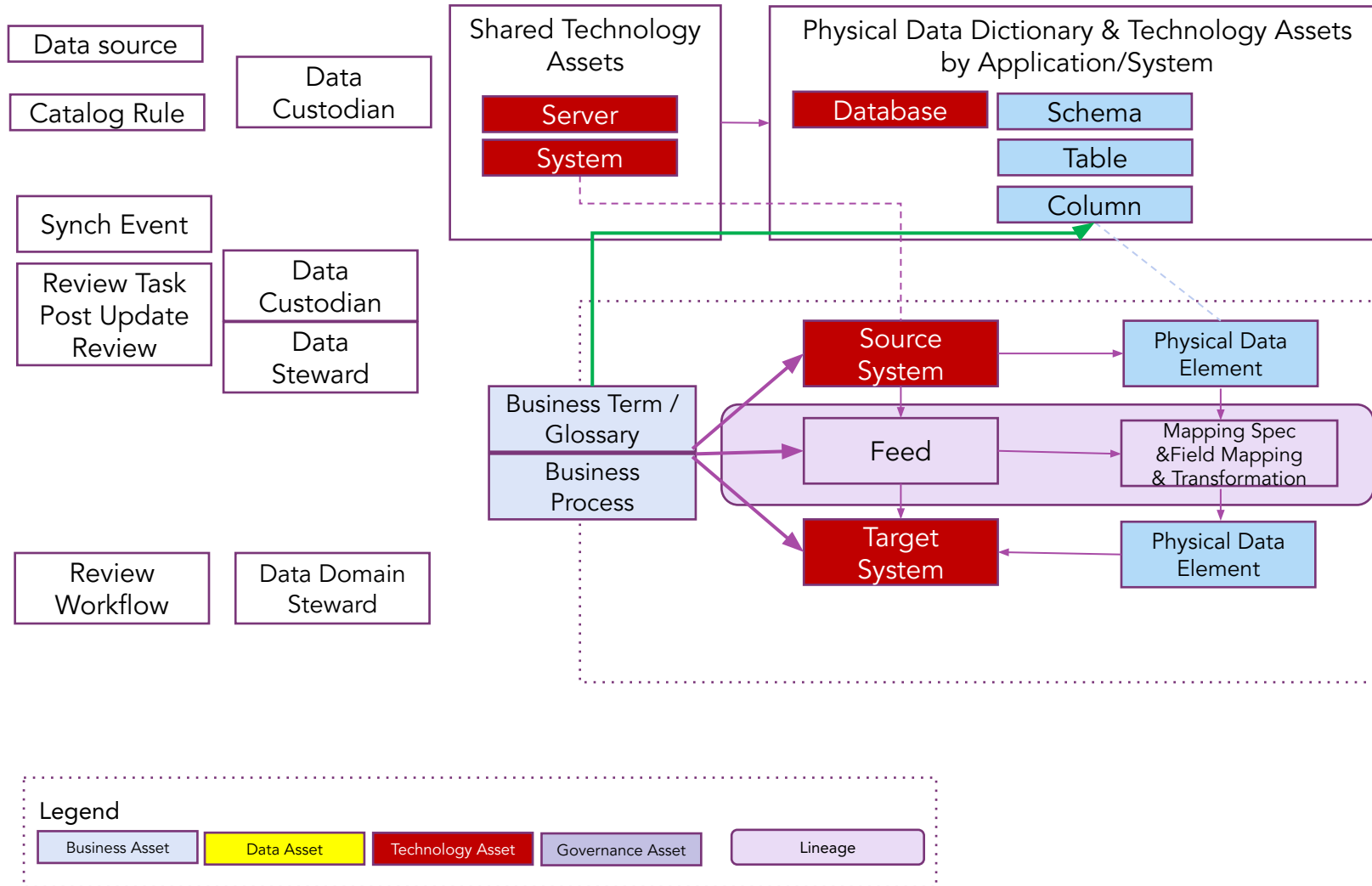
Physical Data Dictionary & Logical Data Dictionary





Metadata

DBMS Metadata Integration Initial Scope



Catalog

Find & Understand



- Platform Features/Highlights
- Manual linking and platform suggestion capability
- Report and DataSet Catalogs
- Shopping for Data & Data Usage
- Advanced suggestion capabilities/integrations; Provisioning

Lineage

Find & Understand

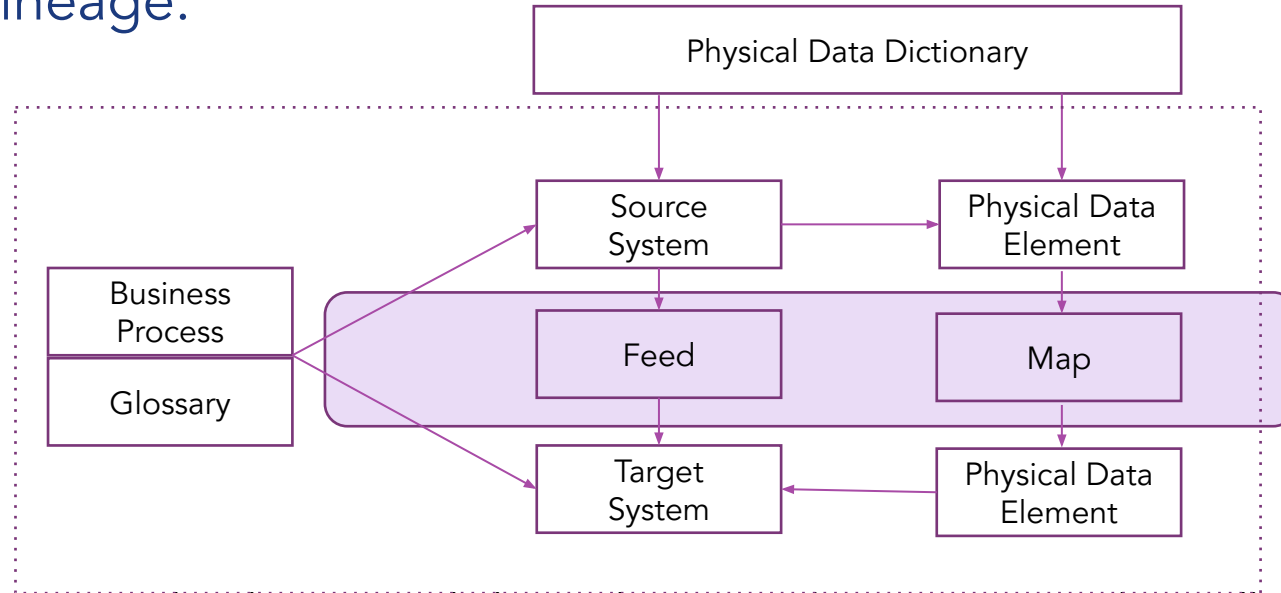


- Platform Features/Highlights
- Sophisticated Template and Ingestion Framework
- Multiple levels – System level and Detailed
- Lineage views – standard views, filtering options, swimlane options, contextual linkages, etc.
...
- DevOps annotation/integration



Lineage

A view on lineage:

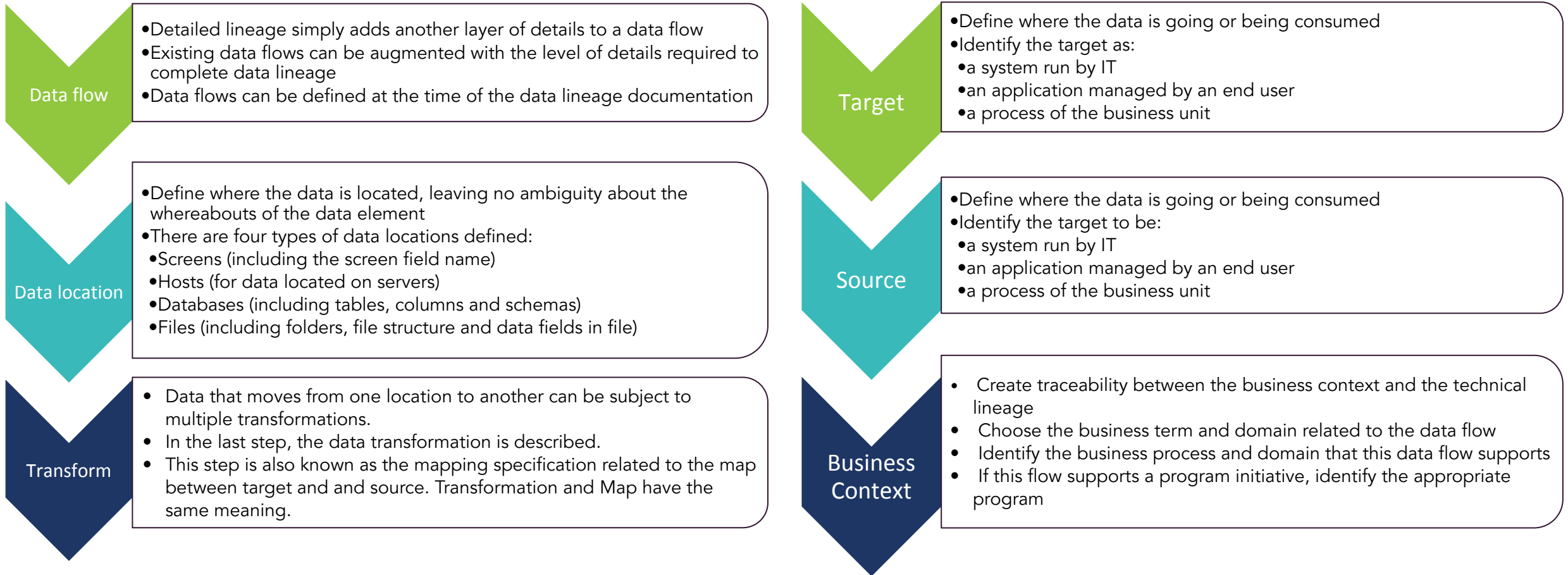


- Lineage is concerned with creating linkages between data assets in order to illustrate the data flow. Data sets are used as a reference to the links.
- Lineage captures:
 - what data is being “fed” between a source and a target
 - the mapping of the data elements between source and target data assets
 - The transformations applied to data elements from the source data asset before it is recorded or used by the target data asset
 - The filters (constraints) applied to source data assets that are subject to transformations, maps or contained in feeds.
- Traceability to other types of metadata is not typically a concern of data lineage. It is an adornment or an augmentation that allows for lineage artefacts to be traced back to other subject areas.



Lineage

Capturing data lineage:





Lineage

Simplifying makes progress:

Data governance platform solutions that work for lineage, risk and control

- Build out a baseline in for Risks and Controls usage.
- Improve the current execution process for capture and recording of Lineage at Domains level.
- Identify gaps, impacts and future solutions on process and development.
- Initiate sustainable self serve model for resource efficiencies with respect to operational support for usage and / or Subject Matter Expertise as needed.
- Write user guidelines and training for stakeholders

Domain

Find & Understand



- Platform Features/Highlights
- Business Process Register
- Policy Register/Link to Enterprise Policies
- Reference Data integrations :
Application and EUC repositories
- Program/Project associations
- Tagging and Collaboration Features

Controls

Trust

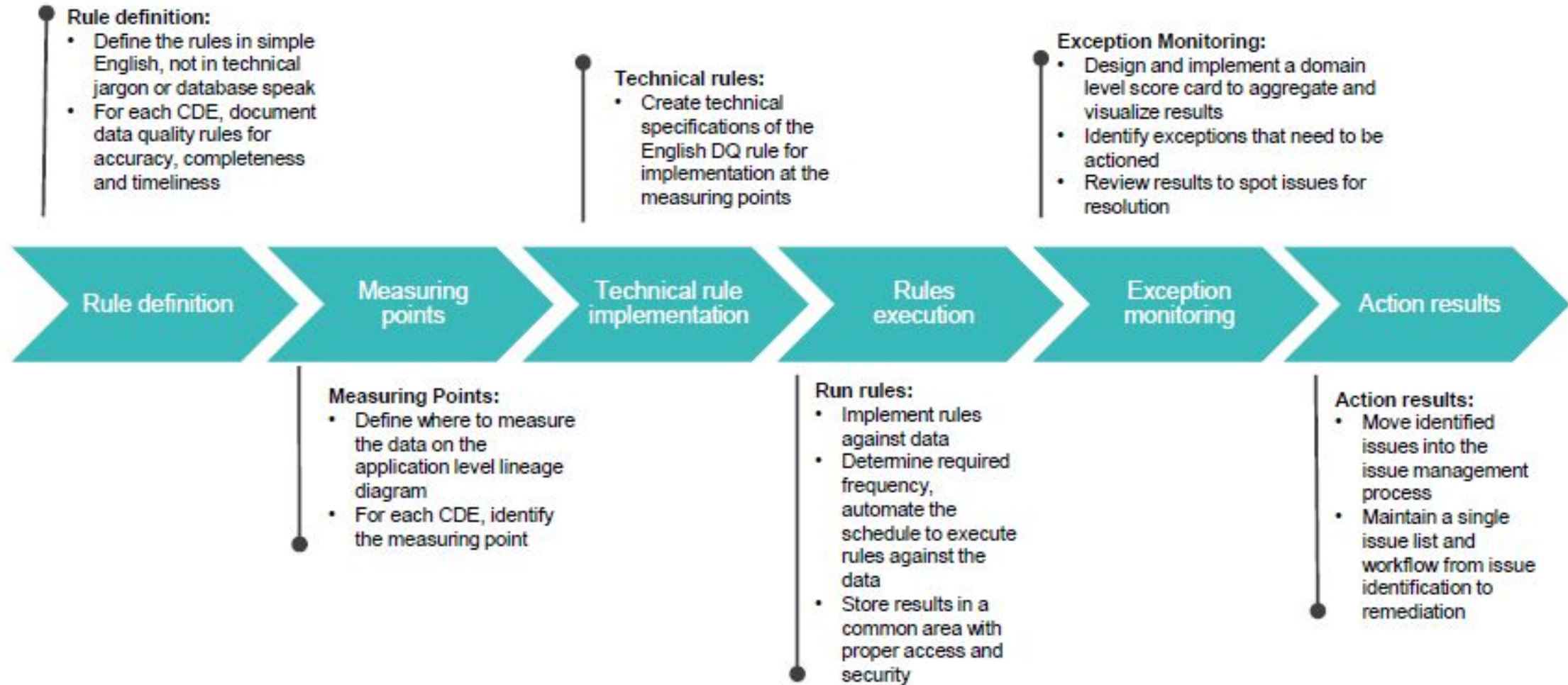


- Platform Features/Highlights
- Risk & Control integration with Risk & Compliance systems
- Data Quality rule loads from QA tools and via workflows
- Results loads from other DQ Engines



Controls

Data Quality Measurement





Controls

Data Risks


Data Risk
The risk of failing to manage information appropriately through its lifecycle due to inadequate processes and controls resulting in legal, regulatory consequences and reputational and financial loss

*With **Petabytes** of Data it needs a holistic understanding of its **data management practices** and **data risk profile** and **appetite** to ensure the proper balance of **risk and reward**.*



Accountability

- ✓ Are we using data up to its **full potential**?
- ✓ Are **data risks** being managed pro-actively?
- ✓ What are key **business & regulatory** requirements around this data?




Data Quality

- ✓ What data is **fit for purpose** and where do we need to focus efforts for fixes?
- ✓ Can we meet the regulatory **threshold** to allow us to use advanced models (e.g. Basel, IFRS9, FRTB)



Data Retention & Destruction

- ✓ Do we really need to hold all that data?
- ✓ Is data being retained **too long** or destroyed **too early**?




Data Sharing & Usage

- ✓ Who are we **sharing** the data with?
- ✓ Are we contractually **protected**?
- ✓ Is data being used **as intended**?



Data Sourcing & Aggregation

- ✓ Where is data coming from?
- ✓ Are data risks being introduced within its flow?
- ✓ Is data being source from **trusted** and authoritative sources



Data Reporting

- ✓ What can this data could do for us?
- ✓ Which are the most **critical** or key reports?
- ✓ Are reports **accurate, consistent** and fit for purpose?

Issue Management & Resolution

Trust



- Platform
Features/Highlights
- Sophisticated workflow with business impact assessment, remediation tasks and collaboration capabilities
- Accompanying reporting and dashboarding

“The Issue Management folks are the detectives of the organization. If you always fancied following Data Operatives’ footsteps, this is the place for you!”

In Conclusion what needs to be done!

Business metadata (e.g. definition, ownership, classification, policies/restrictions) about Data Domains and underlying critical data elements (CDEs) maintenance. Systems lineage including data flows for all CDEs

Full end-to-end data lineage including attribute-level system lineage and business and technical transformation rules for Tier 1 CDEs

Data lineage and metadata information maintenance as changes occur and are reviewed at minimum annually on the Data Governance platform. Data Domain Stewards lead the maintenance of business metadata including domain-specific policy documents, while Data Custodians lead the maintenance of technical metadata

Critical data assigned classification which identifies confidentiality, protection, privacy, access and availability controls or restrictions

Data subject to ongoing assessment of key data risks categories. Preventative and detective controls to address and mitigate identified risks for CDEs

Data Quality defined and monitored for all CDEs to ensure data is fit for purpose. At minimum, data quality measures must cover the following dimensions: Accuracy, Completeness and Timeliness. Other dimensions (e.g. consistency, uniqueness, validity) can be measured where required

Issues related to critical data must be appropriately documented, tracked and remediated. Cross-functional and cross-domain data issues must be logged centrally

Maintain an inventory of approved Authoritative Sources (AS) and Authorized Provisioning Points (APP) which allow data to be shared efficiently with minimal impacts to source Systems of Record (SOR)