What Data and Analytics Leaders Should Know Before Implementing a Data Catalog

Jason Medd

© 2025 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. It consists of the opinions of Gartner's research organization, which should not be construed as statements of fact. While the information contained in this publication has been obtained from sources believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by Gartner's Usage Policy. Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "Guiding Principles on Independence and Objectivity."

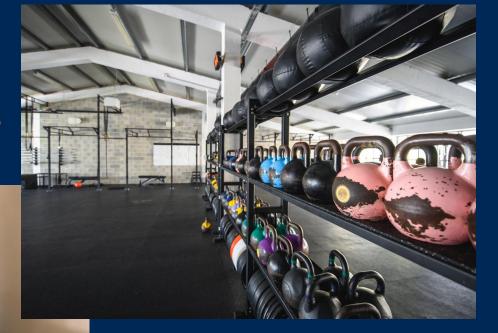


How Many People Made a New Year's Resolution to Exercise More?



You bought the yoga mat.

You had a few sessions at the gym.

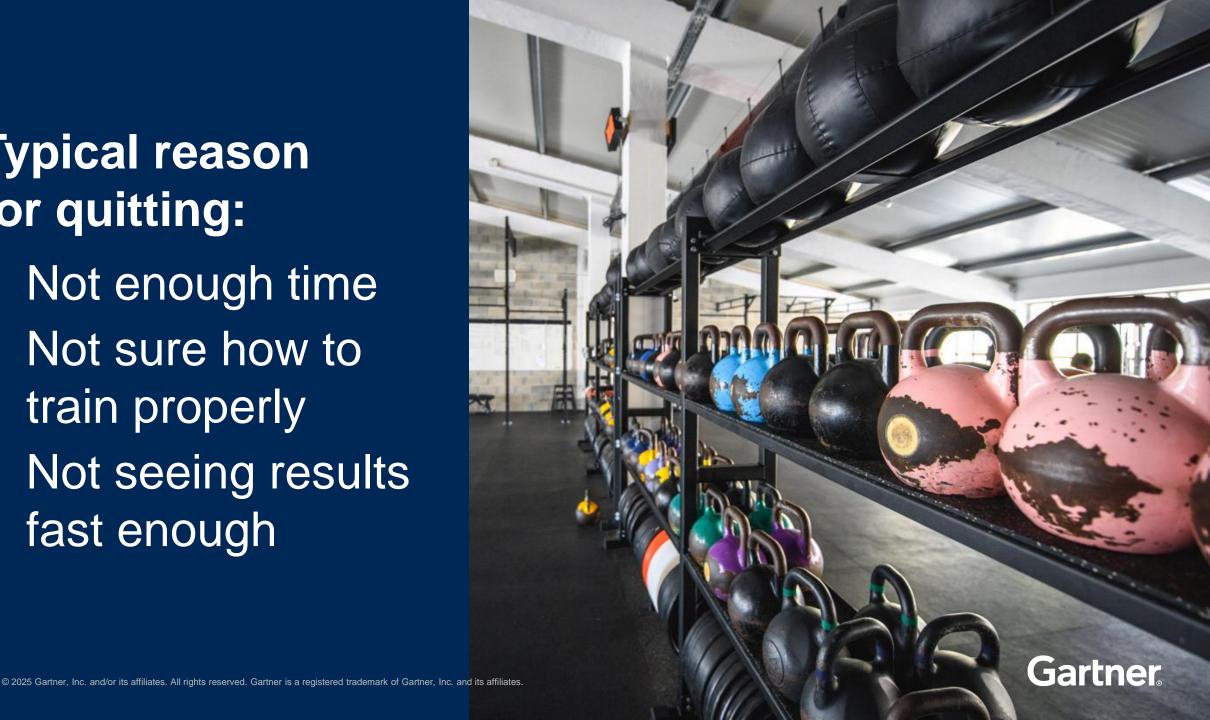


But by March ...



Typical reason for quitting:

- Not enough time
- Not sure how to train properly
- Not seeing results fast enough



Key Issues

- Why adopting a data catalog can be so challenging?

- (?) How data catalogs support data, analytics and Al?

What can you do to improve your chances of success?



Key Issues



Why adopting a data catalog can be so challenging?



How data catalogs support data, analytics and Al?



What can you do to improve your chances of success?



Strategic Issues

- Poor alignment of metadata management needs to business outcomes
- Poor communication on business need and required effort
- Long payback period



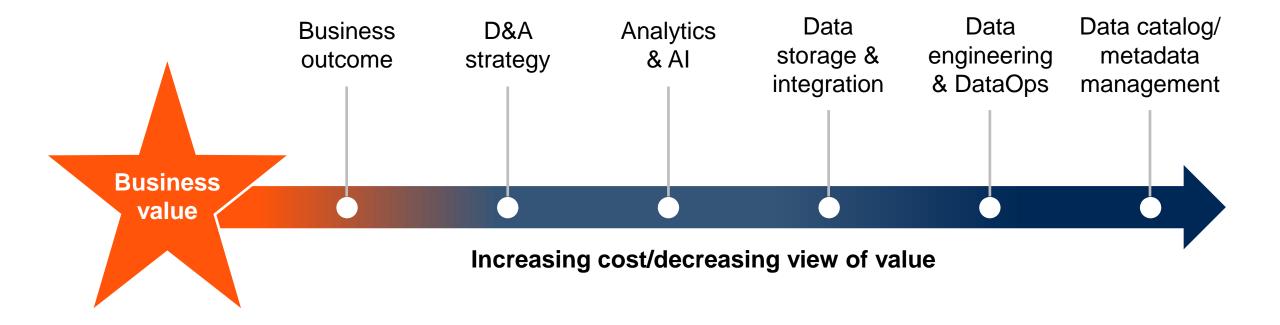


The Role of Metadata in Your Data Strategy

Visible	Locate the needed data		
Accessible	Retrieve the data		
Understandable	Recognize the content, context and applicability		
Connected	Exploit related data elements together		
Trustworthy	Be confident in all aspects of data for decision making		
Interoperable	Have a common representation of data		
Secure	So that data is protected from unauthorized use		



Visibility: Value/Cost Lenses





Communication: What's Needed and Why



Technical

Definitional

- Schemas
- Data types
- Data models



Operational

Descriptive

- ETL/ELT
- Actions on data
- Data lineage
- Performance



Business

Definitional

- Ontology
- Classification
- Tagging
- Mapped business relationships



Social

Descriptive

- Business user knowledge
- User-generated content
- Developer feedback

Low

Context value

High

IT managed

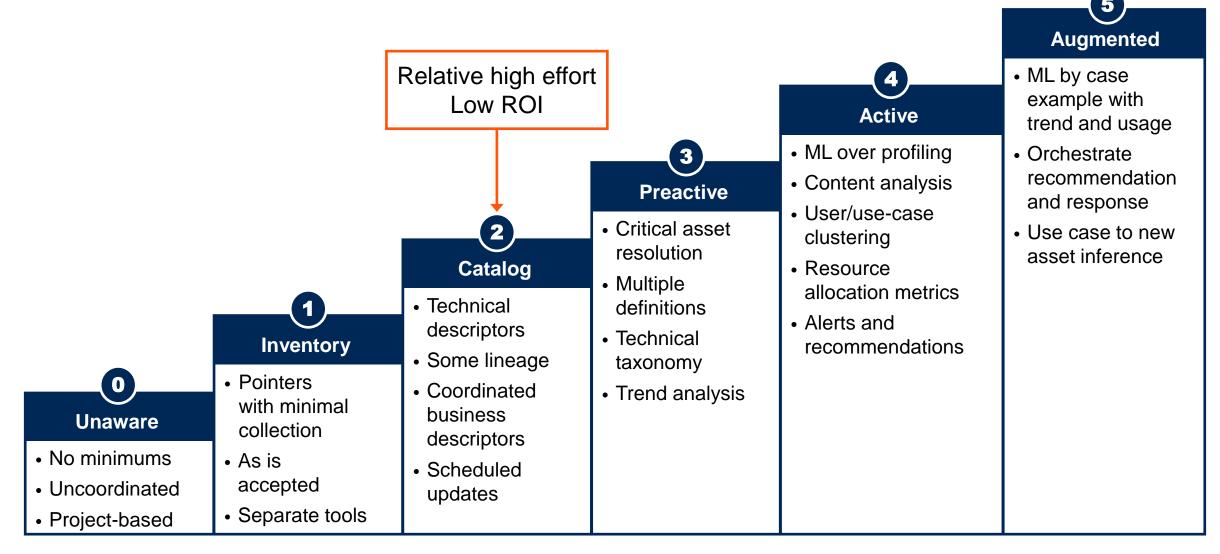
Business involvement

Business managed





Payback Period: Metadata Management Maturity



Source: Gartner

Most Organizations Have Not Invested and Matured Their Metadata Practice to a Stage Where It Can Support Automation in Data Management

Q: Which of the following best describes your metadata management practice maturity?

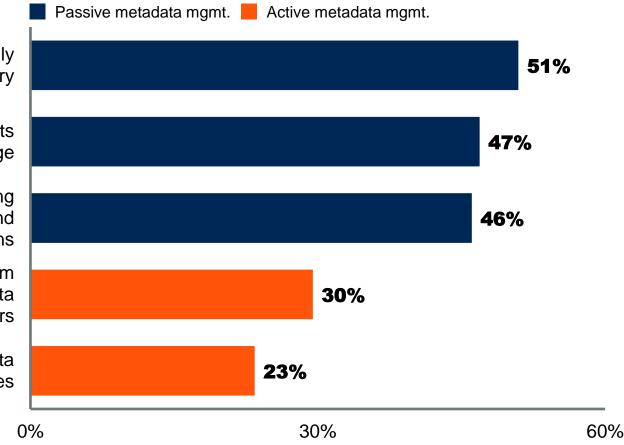
Our metadata practice is early stage and based on a manually maintained data inventory and business glossary

We have implemented an up-to-date data catalog that collects business and technical descriptors and some lineage

We can resolve differences in data definitions under varying business context, we identify critical assets, provide a taxonomy and perform metadata trend analysis such as consumption patterns

We use ML for profiling and classifying data assets, perform metadata analysis and provide alerts and recommendations to data consumers

Our metadata management uses ML by use case to derive data assets, create recommendations and responses



n = 247; data management leaders. Multiple responses allowed.

Source: 2024 Gartner Evolution of Data Management Survey



Tactical Issues

- Skills deficit
- Time deficit
- Poor understanding of metadata requirements
- Wrong starting point
- Adoption/handoff issues





Key Issues



Why adopting a data catalog can be so challenging?



How data catalogs support data, analytics and Al?



What can you do to improve your chances of success?



Use Case 1: Searchability

No. 1: Most described business problem

We can't find our data!

But

- Long payback period
- Self-limiting exercise in some cases



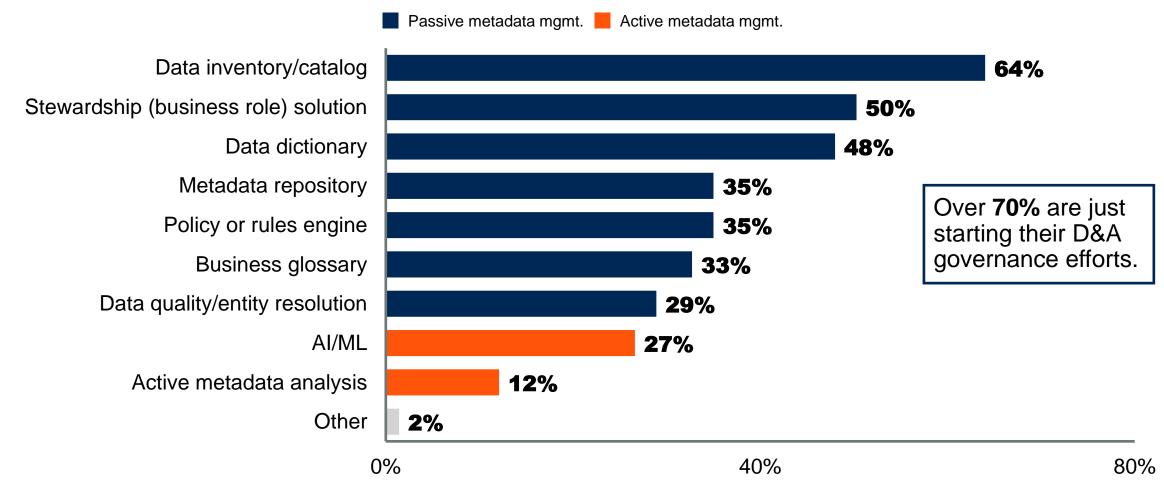
Use Case 2: Data Governance

- Aligns policies, business definitions and critical data
- High success rate when governance is aligned to business outcomes
- Only creates a limited view of data landscape
 - —5% to 10% of structured data



Metadata Capabilities Enable D&A Governance

Q: Which of the following technology capabilities are used to enable governance in your organization?



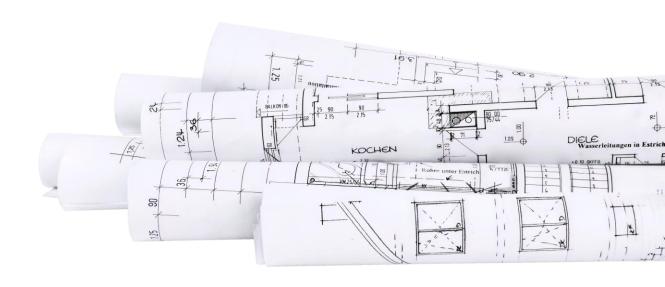
n = 131; D&A governance stakeholders; multiple responses allowed

Source: 2024 Gartner Data & Analytics Governance Survey



Use Case 3: Data Engineering

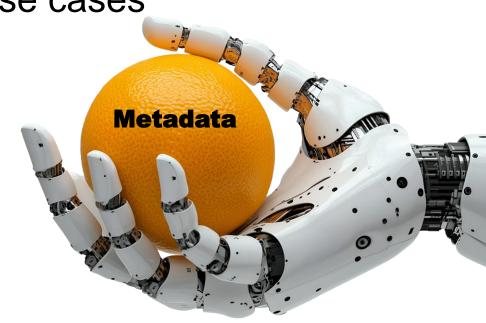
- Supporting impact analysis
- Primarily technical and operational metadata
- Covers critical and noncritical data
 - —Limited business context





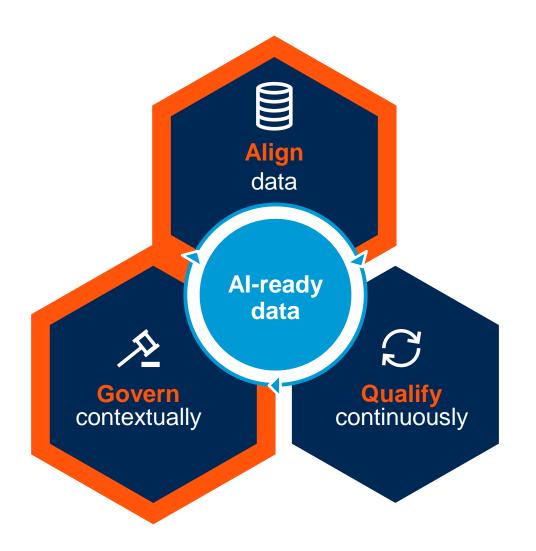
Use Case 4: Al-Ready Data

- Key enabler and pillar for supporting AI projects
- Al is not just data hungry it's metadata hungry
- Composite use case extension of prior use cases
- Same vulnerabilities to adoption
 - —Business outcomes
 - —Visibility
 - -ROI





Use Case 4: Al-Ready Data

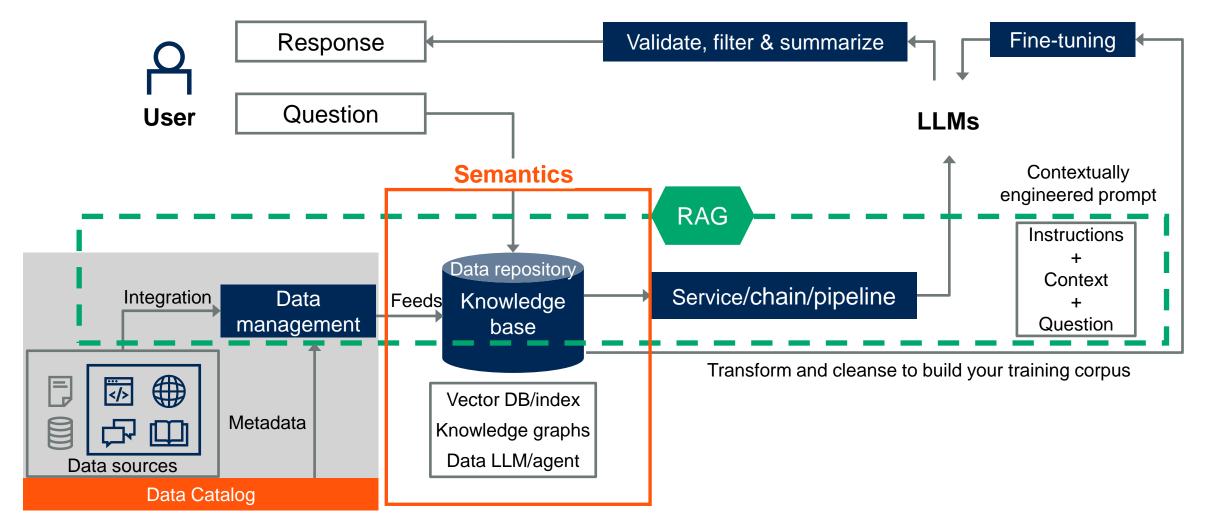


Alignment	Qualification	Governance	
Quantification	Consistency	Data stewardship	
Semantics	assessment		
• Quality	 Validation and verification 	 Inference and derivation 	
Trust and fairness	Operational SLAs	 Regulatory and 	
Diversity	Versioning	compliance	
• Lineage	 Continuous regression testing 	Al standards support	
		Data sharing	
	Observability metrics		

Source: Gartner

LLMs Use Data in a New Way: It Is All About Semantics!









(//) If you are serious about AI, you need to get serious about metadata.



Key Issues



Why adopting a data catalog can be so challenging?



How data catalogs support data, analytics and Al?



What can you do to improve your chances of success?



Strategic Actions

- Communicate a why and articulate expected outcomes
 - —Utilize storytelling
- Stakeholder engagement
 - —Expectation setting on outcomes and effort required
- Education and enablement
 - —Data literacy
 - —Training on usage of the catalog







Remember: Use cases are a "what" not a "why." Do not expect stakeholders to connect the dots by themselves.



Tactical Actions

Gather metadata requirements before buying a catalog

Scoping and approach availability

Assess skill availability

Assess time availability



Metadata Requirements

Create a metadata application profile

- Documents intended usage
- SMEs
- Input rules, vocabulary and formatting standards

Metadata Application Profile							
Element Name	Mapped to Dublin Core	Obligation	Vocabulary/Encoding Scheme	Input Guidelines	Examples		
Author	Creator	Required	Library of Congress Name Authority File (LCNAF)	Enter the name(s) of the creator(s) of the object. Construct names according to Library of Congress Name Authority File.	Hoff, Syd, 1912- / Skofield, James		
Title	Title	Required		Transcribe title from the resource itself, such as the book title or object name, using the same punctuation that appears on source.	Jurassic poop		
Illustrator	Creator	Required if available	Library of Congress Name Authority File (LCNAF)	Enter the name(s) of the creator(s) of the object. Construct names according to Library of Congress Name Authority File.	Teague, Mark		
Publisher	Publisher	Required		Enter name(s) or institution of publisher of the object.	HarperCollins / Dial Books for Young Readers		
Copyright Year	Rights	Required		Enter copyright date as listed within copyright information of source.	1993 / 2006		

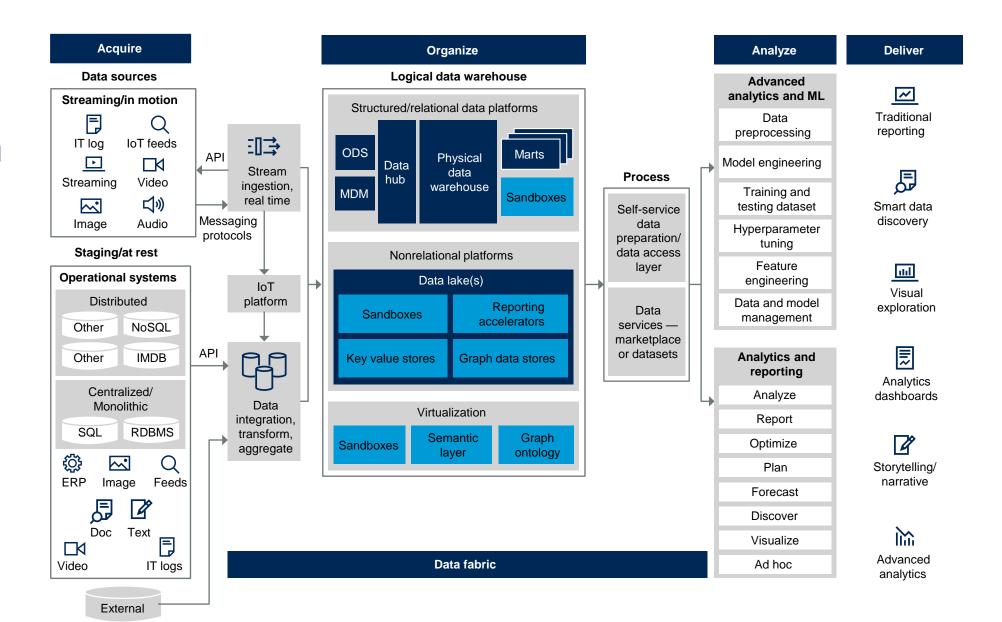


Scoping and Approach

End-to-end architecture for data management, analytics and DSML

Manage and govern:

information governance (metadata management, data quality, data modeling and master data management), data management (data admin, security, privacy and identity) and organization (people)



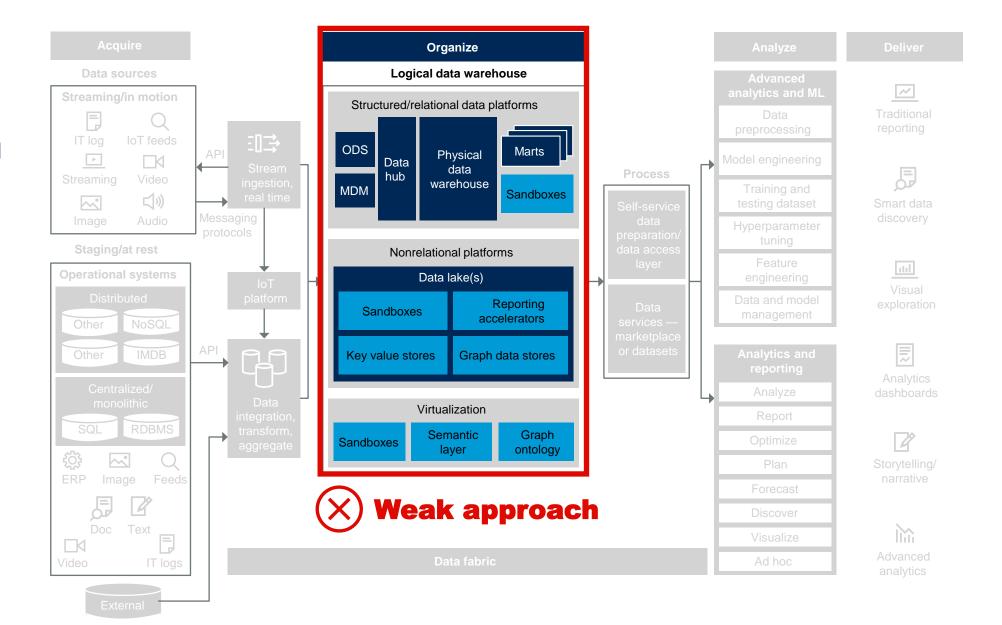


Scoping and Approach

End-to-end architecture for data management, analytics and DSML

Manage and govern:

information governance (metadata management, data quality, data modeling and master data management), data management (data admin, security, privacy and identity) and organization (people)



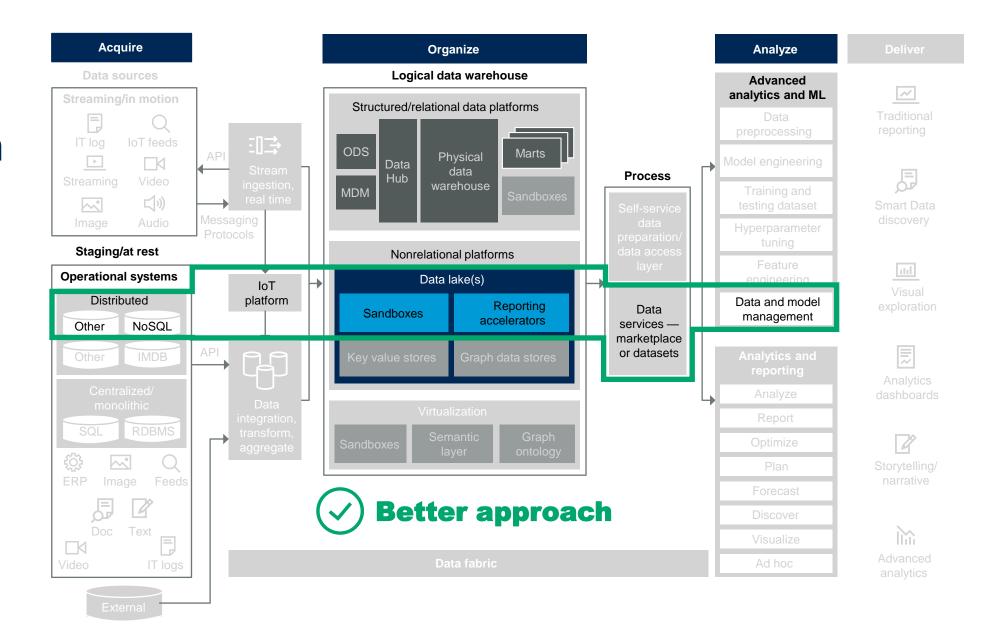


Scoping and Approach

End-to-end architecture for data management, analytics and DSML

Manage and govern:

information governance (metadata management, data quality, data modeling and master data management), data management (data admin, security, privacy and identity) and organization (people)





Skills

Cataloging is a skill

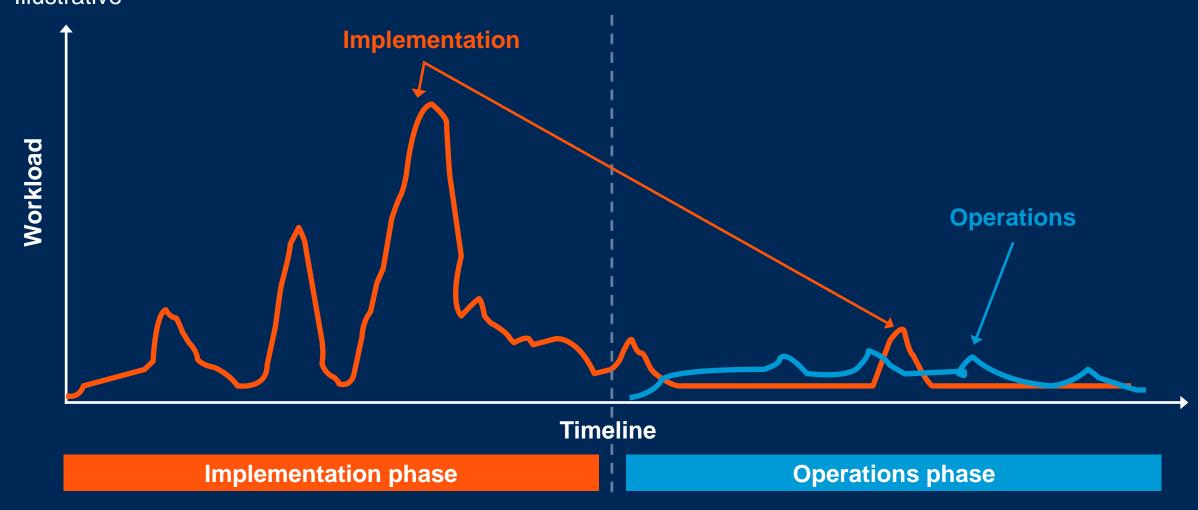
- Taxonomies
- Ontologies
- Controlled vocabularies
- Defining and enforcing standards
- Enabling users to effectively use a catalog





Time

Workload Based on Implementation and Maintenance of a Data Catalog Illustrative





Roadmap for Deploying an Enterprise Data Catalog

Preparation (3 months)

Implementation phase (6 months)

with the vendor (7 to 24 months)

Standardization phase (beyond 2 years)

Fact finding

Inward focus: building your business case, setting up governance framework and finding answers to:

- Where is data (apps)?
- o Who knows?
- o Who owns?
- o Who uses?
- Outward focus:
 market trends on
 metadata management,
 transformational use
 cases, top vendors.

Prototyping

- Vendor assessment: technical support during POC, reference customers, costs, vendor alignment with your future strategy like hybrid/multicloud.
- Tool assessment: product core capabilities like connecting to data sources and creating an inventory of data assets, profiling, tagging, searching, usage metrics, supporting both passive and active metadata.
- People assessment: ease of finding data, understanding context, maintaining the inventory and governance.
 Should involve varied personas from business and IT.

i**on:** un-and

Pilot

- Installation: up-andrunning live and test platforms with necessary access controls.
- Connectivity: first set of data sources connected to the DC tool.
- Inventorying data assets, classifications.
- Business glossary, lineage enhancements to the inventory.
- Train the trainer (mandate this to the vendors as a prerequisite in your statement of work).

Show-and-tell

Expanding

- clinics to educate users about the tool and process.

 Spontaneous
- Spontaneous
 addition of new
 users: efforts
 focused on building lata communities.
- Addition of new data sources.
- Rewarding top contributors.
- Measuring productivity, selfservice usage.
- Set up a competency center.

Data literacy • Community

- efforts: harvest communal data knowledge, refine governance framework with standards to drive consistency.
- Metrics to
 measure data
 literacy such as
 the volume of
 communal data
 knowledge
 harvested and
 productivity of
 business users.

Data automation

- DC becomes the core of your ecosystem by standardizing DC usage for all data producers and consumers across organization as an enterprise hub for systems of record for data.
- The contributors and consumers begin to blend and reach a plateau stage; which is a mark of increased data literacy within the organization.



Recommendations

- Supplement people and process of a data catalog project with technology. Do not offer technology as a replacement for people and process.
- Build user engagement for a data catalog by using storytelling to describe both what can be possible thought-effective metadata management, and to describe the challenges and pitfalls along the way.
- Advocate the need for appropriate skills in managing standards and procedures in operationalizing the data catalog by identifying obvious gaps in knowledge and enthusiasm from the steward community for the job.



Recommended Gartner Research

To learn more about access to Gartner research, expert analyst insight, and peer communities, contact your Gartner representative or click on "Become A Client" on gartner.com to speak with one of our specialists.

- How D&A Leaders Can Use Metadata for Better Business Outcomes
 Thornton Craig
- State of Metadata Management: Aggressively Pursue Metadata to Enable Al and Generative Al Mark Beyer and Guido De Simoni
- Successful Generative Al Projects Require Better Metadata
 Management
 Mark Beyer, Roxane Edjlali, Ehtisham Zaidi, Melody Chien and Others
- What Data Architects Need to Know About Data Quality Jason Medd

