# Welcome to the Architect Workshop Metadata Driven Data Management



### Today's Presenters



Data & Analytics Deloitte

**Principal Architect** Informatica

Associate Director IPS Informatica

Solution Architect Informatica



### Agenda



Overview, Trends and Requirements of a Metadata Driven Data Management Approach

Srikant Kanthadai, Deloitte



Solutions and Reference Architectures enabling Metadata Driven Data Management

Manuel Gomez, Informatica



Organizational Perspectives -Bringing Metadata Architectures to Life

Trevor Hodges & Claire Rolland, Informatica



**Q&A** Session

All Attendees



### Deloitte.

### Emerging Trends in Metadata Management

00100

8181

000

1000 H

1

- BBBBABAA

1110101001

1 0 1

GOODA.

Srikant Kanthadai, Deloitte MCS UK

September 2021

### Agenda

- Metadata Management Concepts
- Key Capabilities
- Application of the Solution
- Emerging Trends
  - The era of big data
  - Business focussed metadata initiatives
  - Prioritisation of data quality
  - Metadata, master data and data privacy
  - Operationalising Metadata Management
- Maturity Framework

#### Metadata Management concepts

#### Key Metadata Management concepts to ensure a successful, business-oriented approach

Metadata management is about management of your data and information assets to address use cases such as data governance and analytics. It is used as a reference for business-oriented and technical projects and builds the foundations for describing, inventorying and understanding data.



### Key capabilities

#### A simple model for unlocking the value of data



#### Application of solution How your problems map to capabilities

Current data-related challenges facing the clients resolve around few key concepts: **customer data, black box, lack of single view of overall data, lack of definitions and governance, lack of mapping and quality of data** 



© 2020 Deloitte MCS Limited. All rights reserved.

#### Emerging Trends – the era of big data

- Shift towards *data centricity & data monetisation* looking to *identify*, *classify & gain insights* from their data assets.
- Insight go beyond the basics of mapping & understanding metadata to how it is *mined, used & managed*
- Shift seen not only for metadata management but also in the governance of data *creating a trusted data platform*

#### Emerging Trends – business focussed metadata initiatives

- Focus on added *business context* to technical metadata
- Help provide clarity on how to *effectively leverage* your *data asset*
- Confirm data is *trusted & fit for purpose*

### Emerging Trends – business focussed metadata initiatives

- Focus on added business context to technical metadata
- Help provide clarity on how to effectively leverage your data asset
- Confirm data is trusted & fit for purpose

### Emerging Trends – prioritisation of data quality

- Link data quality initiatives to metadata initiatives
- Use metadata to *measure data quality*
- Documenting *data quality standard in a catalogue* and using the same to enhance data quality
- Establish *trust in critical data assets* that impact business goals & objectives

#### Emerging Trends – metadata, master data & data privacy

- Regulations such as GDPR, TRIM, and CCPA, the very definition of "*data privacy*" *is a moving target* need to protect against unauthorized access, usage or dissemination
- Using metadata to *identify data stakeholders, PII attributes & process to protect* the data (confined data, masked data, restricted access etc.)
- Metadata always plays a vital role in *managing master data*, *ensuring it's accuracy and consistency* across systems. This will ensure consistent use of the data attributes across source/ consuming systems and also help capture lineage

#### Emerging Trends – operationalising metadata management

- Metadata can *become stale* very quickly as data source change
- Need for constant *proactive* metadata refresh and monitoring over time
- *Automating metadata* to become *living source of information* -metadata collection, quality metrics & crowdsourcing business context
- Use AI techniques to *automate mapping of metadata* to business defintions

### Maturity Framework

#### Your current position on the maturity framework of data assets

Maturity Level		1	2	3	4	5
		Initial	Managed	Tool-based	Optimized	Automated
Discovery		Data assets are found by chance or personal network	Data assets are listed in documents, wikis or spreadsheets	Data assets can be searched in a central catalogue based on a dedicated tool	The catalogue can display sample data entries and summary statistics	A recommender system for data assets is in operation
Trust		An evaluation of the data properties is hardly possible	A contact person, e.g. the data owner is listed for each data asset	Comments and tags indicate the properties of the data asset	A lineage and impact graph shows where the data comes from and how it is used	Quality metrics are available for all data assets
Provision		No metadata is provided for central use	Metadata is provided on request by the data owner	The data catalogue recognises and labels standardised data types, e.g. account numbers	Lineage detection within the data platform is highly automated	Machine learning algorithms support the provision of metadata across platforms
Collaboration	<u>I</u>	There is no cross- functional cooperation	Templates help to structure the generation of business metadata	A business taxonomy simplifies and standardises the capturing of business metadata	Creation of business metadata is gamified or facilitated by nudges	Significant efforts in the capture of business metadata are rewarded by a financial bonus
Governance		Metadata related roles and responsibilities are unclear	Metadata related roles and responsibilities are assigned	A central metadata managements team coordinates the work	All metadata governance processes are supported by workflows	Data access and approval processes are largely automated

Your position on the framework is a result of the challenges you are currently facing:

- Data dictionary
- Logical Modelling
- Physical Modelling
- Data Catalogue
- Metadata Management
- Master Data
- Reference Data
- Asset Inventory
- Quality of Data

### **Deloitte.**

This publication has been written in general terms and we recommend that you obtain professional advice before acting or refraining from action on any of the contents of this publication. Deloitte MCS Limited accepts no liability for any loss occasioned to any person acting or refraining from action as a result of any material in this publication.

Deloitte MCS Limited is registered in England and Wales with registered number 03311052 and its registered office at Hill House, 1 Little New Street, London, EC4A 3TR, United Kingdom.

Deloitte MCS Limited is a subsidiary of Deloitte LLP, which is the United Kingdom affiliate of Deloitte NSE LLP, a member firm of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"). DTTL and each of its member firms are legally separate and independent entities. DTTL and Deloitte NSE LLP do not provide services to clients. Please see www.deloitte.com/about to learn more about our global network of member firms.

© 2020 Deloitte MCS Limited. All rights reserved.

Solutions and Reference Architectures enabling Metadata Driven Data Management

Manuel Gómez, Principal Architect



### Becoming Metadata-Driven is a Journey Architecture to enable Metadata-Driven Management



### Metadata Management Reference Architecture







🔶 Informatica















### Metadata Management Reference Architecture



Organizational Perspectives - Bringing Metadata Architectures to Life

Trevor Hodges, Associate Director EMEA IPS

Claire Rolland, Solution Architect - Strategic Accounts



### INFA Metadata Management Capability Framework



### Your path to success with Metadata Management



#### Your goal is to maximise the reach & impact of metadata



### Two basic program Strategies (or a mix of the two!)

**Retail Bank** 

Global Consumer Packaged Goods

#### Build from outside in

 A program strategy based on your customer's needs



- A program strategy based on what you think what's best for your customers...
- Or you think you know what's best for your customers



### Our retail banking customer started with 'why'



#### They defined their program strategy that:

- Aligned and supported its corporate vision
- Focused on delivering business impact with use cases linked to business drivers
- Addressed user pain points
- Communicated program value and long-term direction

# Their program demonstrated value - strategy-based Vision, Business Drivers, Use Cases, and Pain Points

2

່ 3

#### **Corporate Vision**

 Aspire to provide the best retail banking service for our customers

#### **Business Drivers**

- Drivers
  - Increasing regulation (external)
  - Revenue growth (internal)
- Business initiatives/tactics
  - 360 view of customers to Support online customer selfservice
  - Migration: DWH to cloud

#### **Required Metadata capability use cases**

- End to end lineage of key data elements
- Inventory of data assets
- Metadata Architecture

#### **Challenges and Pain Points**

- Creating data lineage compliance report of key data elements is two
  months of manual effort each year
- Data quality is questionable, not sure who is responsible
- No common business definitions causing data consistencies
- Data analysts and data scientists spends >75% of other time finding trusted, relevant data sets for analytics

#### **Data Catalog Program Strategy**

• Democratize data asset knowledge and usage through a central data catalog to support cross-functional collaboration, self-service analytics, IT modernization, and Data governance



### Delivering a 'google like' view of the bank

learn from pilot, deliver quick wins, and expand



### They trained and supported their users

1

#### Created Training Plan by User Role

- Created training curriculum by user roles
- Trained its support staff first; they were internal functional and technical EDC experts
- Leveraged Informatica University ondemand training, IPS Adoption Services
- Customized training for its users

#### 2

## Created Training Content and Videos

- Engaged its subject matter experts to define hands-on training content
- Compared and contrasted the current process versus using EDC to demonstrate the benefits of using EDC
- Communicated the bigger picture of its program strategy and set user expectation
- Created short 3-4 minute videos as a how-to reference for performing common tasks

#### Leveraged Business Champions to Train and Drive User Adoption

3

- Created business ownership by leveraging its champions as influencers to drive adoption in their teams
- Leveraged its enablement experts
- Created training image on cloud to quickly spin-up training images and future training upgrades



### They measured their results and asked for feedback

#### Monitor and Measure

- Monitored usage with metrics including user logins, searches, reviews, time-saved, new asset usage, data set discovery, new datasets created
- Conducted user satisfaction
   surveys

#### Engage

- Explained the benefits
- Encouraged collaboration on data with review and ratings on datasets.

Feedback
<ul> <li>Conducted follow-up user workshops</li> </ul>
<ul> <li>Discussed what works, what can be improved</li> </ul>
<ul> <li>Shared tips/tricks</li> </ul>

Identified what need to be fixed



### They then refined and expanded the project

- Developed a 3-year roadmap aligned to corporate/compliance needs
- Incrementally added more users to the catalog
- Expanded to new use cases privacy management
- Cultivated data-oriented community around the data catalog
- Created custom applications to leverage its enterprise data knowledge captured in the catalog - REST APIs
- Increasing automation leveraging AI capabilities
- Apply Custom Scanner Framework



### Our Global CPG approached from the inside-out

#### Data initiatives take time, and require continued effort long after the original project is completed. Expectation management is key so the community has a long-term view of what is expected of them, and where the value is 1. Initiation 3. Completion Lack of engagement from Failed to demonstrate value the start - no real buy-in, Poor transition from project 2 sponsorship, budget, Once the symptoms were relieved, the support etc. effort stopped 2. Execution 4. Growth • No true engagement from community • No plan for what happened Didn't relate to business needs next Required significant learning effort Enthusiasm waned, declined into irrelevance

#### Open to organisational disruption

Poor expectation management

A single point of failure, not supported broadly enough, so that the inevitable changes in the organisation upset the progress. The programme and solution design was also inflexible, and could not cope with change.



### Summary

#### Formulate Strategy

Align corporate business direction Plan strategically, start with users in mind Outside in thinking

#### Start with a Pilot

Keep first rollout simple, focused; quick wins

Make catalog relevant to users with business context information. Start with 'why'

#### Train Users

Train and show users how data catalog helps them produce results

Engage early adopters as your champions

#### **Refine and Expand**

Conduct feedback workshops, monitor usage

Communicate success and expand program







NARRAN

AM S RU S TARATA

REFERE VOIS II CO

HELL

SIE (FIS FIS