Data Asset Catalog Service (DACS) Overview

State Government GIS Users Committee

August 23, 2018

Agenda

- Problem Statement
- What is Metadata Management?
- Metadata Management Process
- DACS Service Platform
- Metadata Capture and Taxonomy
- DACS Service Governance Process
- Value Propositions
- Q&A

Problem Statement

- Many vertical/siloed applications across the agency/State enterprise
- No single/centralized metadata repository for agency/State data assets
- Lots of data assets in many different places and in many different forms
- Lack of SME's familiar with all of the data assets that are available
- No standardized taxonomy for collecting and organizing agency/State data assets
- Slow responsiveness to urgent data related to analytical needs/challenges
- Limited business/technical documentation and/or metadata available in many cases for agency/State data assets
- Some data assets being leveraged across multiple applications but this is the exception
- Some data assets are very well understood and being managed or leveraged well while many others are not
- Lots of data duplication, data quality, data usage, and other inefficiencies as a result of the above

A very typical set of conditions and circumstances for a State enterprise such as ours

NC DIT

What is Metadata Management?

Metadata (National Information Standards Organization-NISO)

Structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. In simplest terms, metadata is "data about data".

Metadata Management (Gartner)

The business discipline for managing the metadata about the information assets of the organization.

- Many different types of metadata can be managed in a variety of ways with business and technical metadata being most important to us
- Metadata management is critical for the efficient and effective reuse of data assets for business-led data analytics and data integration initiatives
- A critical capability underpinning any information management program/strategy

Are you currently doing metadata management in any form or fashion?

Metadata Management Process

Business vs. Technical Metadata

- The following are examples of types of business & technical metadata.

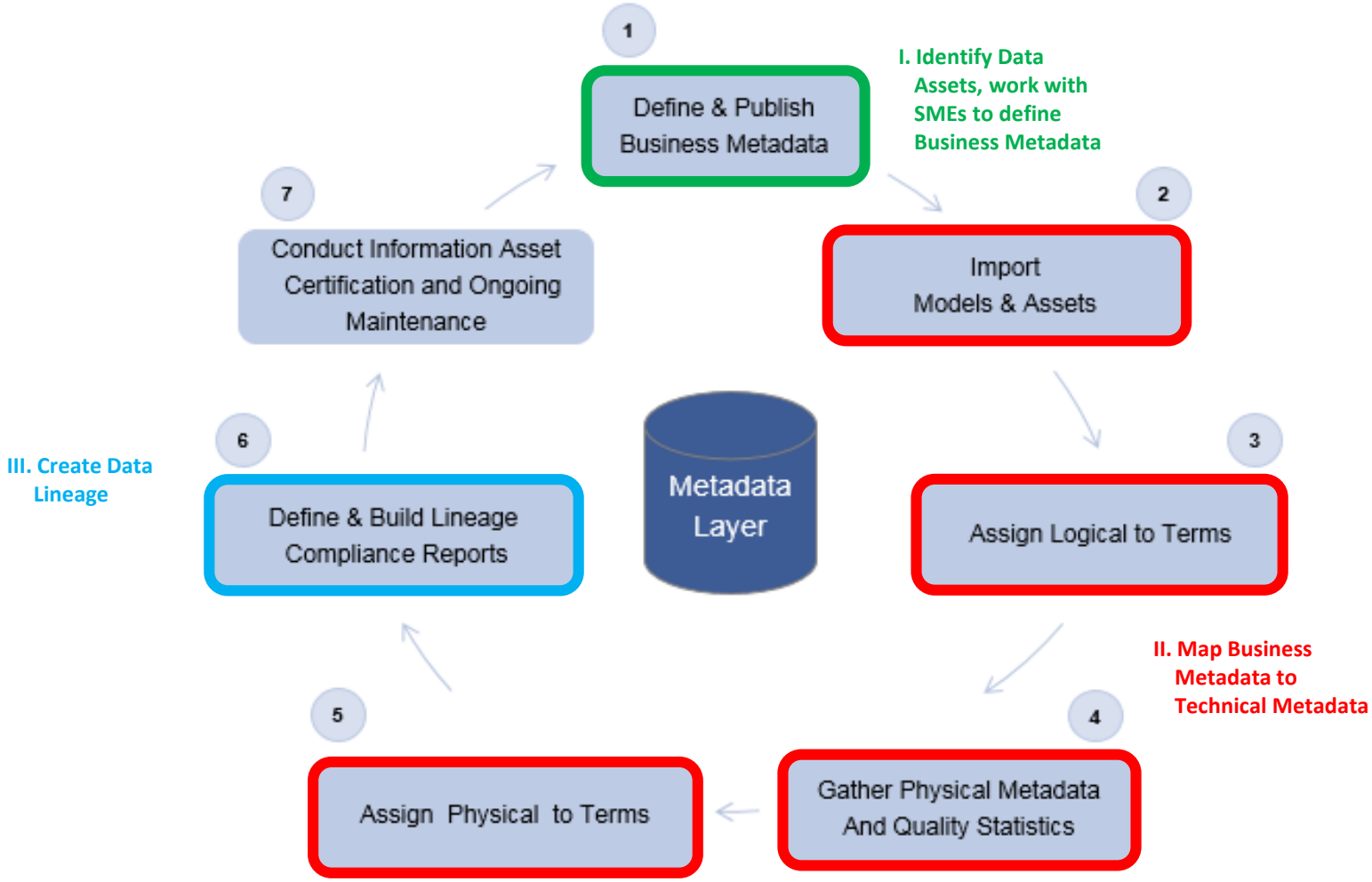
"A customer is a person or organization who purchases a product or service from..."

"CUST_LNM is VARCHAR(30) on the Oracle database CustDB1.

Business Metadata	Technical Metadata
• Definitions & Glossary • Data Steward • Organization • Privacy Level • Security Level • Acronyms & Abbreviations • Business Rules • Etc.	• Column structure of a database table • Data Type & Length (e.g. VARCHAR(20)) • Domains • Standard abbreviations (e.g. CUSTOMER -> CUST) • Nullability • Keys (primary, foreign, alternate, etc.) • Validation Rules • Data Movement Rules • Permissions • Etc.

Global Data Strategy, Ltd.

Metadata Management Process



Achieving business outcomes by building a metadata repository for information governance is an iterative process

Graphic Courtesy of IBM

Metadata Capture and Taxonomy

Example : What types of metadata is the GDAC collecting for the State's assets?

The GDAC has defined a taxonomy for categorizing/organizing all of the assets

1. **Business Metadata**
   - ***Business Domain*** – business taxonomy for classifying or categorizing data file, e.g. health, vehicle, adult correction, etc.
   - ***Short Description*** – to understand it's purpose and use
   - ***Data Source*** – state agency that is creating the data
   - ***Source Type*** – To whom does the data belong? e.g. federal, state, etc.
   - ***Data Classification*** – level of security around data. **NC Data Classification and Handling Policy**
   - ***Data Compliance*** – subject to which data privacy/protection e.g. HIPAA, FERPA, etc.
   - ***Applications*** – inventory of which applications are using any given data asset
   - ***Data Stewards/Data Custodians*** – the names and contact info for the person primarily responsible for maintaining the metadata for any given data asset catalog entry
   - ***Agency/Department/Division Data Stewards*** – the name and contact info for the person primarily responsible for the metadata at the source for any given data asset which may or may not be the same name as listed above
   - ***Status*** – is the data asset active, inactive, etc.
   - ***Frequency of update*** – how often does the metadata for any given asset get updated (e.g. daily, weekly, monthly, etc.)

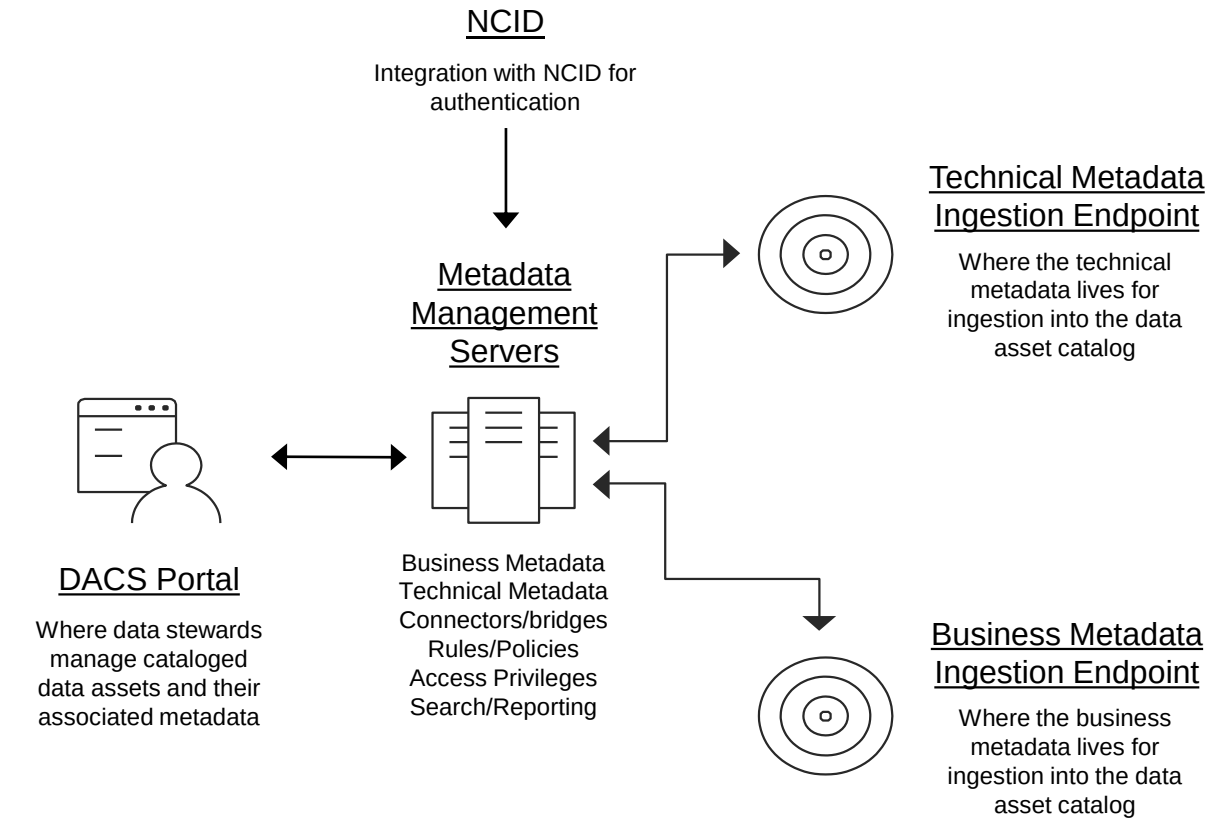
Metadata Capture and Taxonomy (continued...)

2. **Technical Metadata**
   - ***Data hosting environment*** – a list of server names, OS versions, and their metadata
   - ***Data storage*** – database, schema, table, and column names
   - ***Database size and storage information*** – how much storage space is allocated and being used
   - ***DB platform and version*** – what DB platform and version does the data resides on (e.g. Oracle 11G, DB2 10.5, etc.)

There is a “standardized” set of metadata attributes as well as an infinite number of custom metadata attributes that can be created and captured for any given data asset

DACS Service Overview

- A centralized metadata repository for State data assets
- Ability to mechanically ingest existing business and technical metadata from any database or data source
- Facilitates holistic viewing and reuse of State data assets
- Standardized/shared vocabulary for data across the State enterprise
- Extensive search and reporting capabilities on cataloged data assets
- Decentralized/federated management of data assets by data owners
- Less dependency on specific SME personnel for data-related decision making/reuse
- A key foundation that is needed to support enterprise data governance/management and Data-as-a-Service (DaaS)
- Rapid on-boarding/ramp-up using an existing multi-tenant, shared service type platform
- No cost to State agencies for use of the DACS



You can't do enterprise data governance without enterprise metadata management

DACS Service Governance Process

- A customer-led governance council and process
- Current council membership includes reps from DOT, CGIA, GDAC-Solutions Development, and Commerce
- Council chair is elected by council membership for a (6) month tenure (or longer)
- Reps from customer agencies hold voting member seats on the council
- Governance council meets monthly or quarterly on a recurring basis
- Working with the council chair, the DACS service owner enables and facilitates the governance process as a non-voting council member
- Consensus decision making is preferred by the council but when not possible for any reason, the majority vote by the membership rules

A customer-led governance process is essential in order to ensure that the service is meeting the needs of the people that are using it

Value Propositions (or "Why is DACS worth it?")

- A centralized metadata repository for all agency/State data assets
- Ability to mechanically ingest/import existing business and technical metadata from any database or data source
- Facilitates holistic viewing and reuse of agency/State data assets while enabling a more rapid response to ad-hoc data integration and analysis needs
- Standardized/shared vocabulary for data across the agency/State enterprise
- Extensive search and reporting capabilities on cataloged data assets
- Decentralized/federated management of data assets by data owners
- Less dependency on specific SME personnel for data related decision making/reuse
- A key foundation that is needed to support enterprise data governance, Data as a Service (DaaS), and Master Data Management (MDM)
- Rapid on-boarding/ramp-up using an existing multi-tenant, shared service type platform
- No cost to State agencies for use of the DACS

A best-in-class platform for capturing and managing business and technical metadata for the State's data assets at an enterprise level

NC DIT

Q & A

Babita Savitsky
DACS Service Owner, Technical Lead, and Governance Council support
babita.savitsky@nc.gov
919-754-6184

David Giordano
DACS Governance Council Chair
david.giordano@nc.gov
919-754-6585

For more information: https://it.nc.gov/dacs