





CPI-DASG-
DISM-DMDG




 Pages

SPACE SHORTCUTS

 DMDG Validation En...

 DMDG Production E...

 DMDG SNOW Help ...

Space Tools

[Overview](#) [Content Tools](#) [Tasks](#)

[Reorder Pages](#) [Undefined Pages](#) [Attachments](#) [Export](#) [RSS Feeds](#)

▼ [DMDG](#) 

[DMDG Calendar](#)

▶ [DMDG Application Onboarding](#) 

▶ [DMDG Documentation Library](#)

▶ [DMDG Internal Documentation](#)

▶ [DMDG Project Management](#)

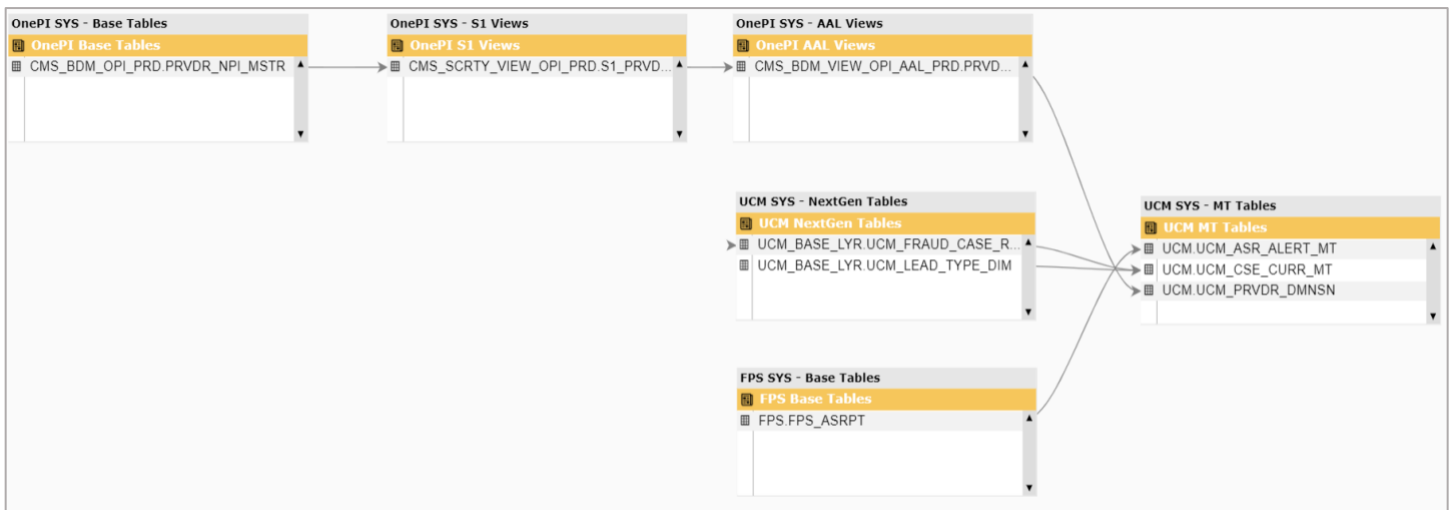
About CMS-DASG-DMDG Confluence wiki

The DMDG team created the CPI-DASG-DMG Confluence wiki to share, publish, and collaboratively edit and update technical documentation, training and onboarding materials, and known DMDG Erwin DI defects.

The wiki serves as a one-stop resource for internal and external stakeholders and end-users to read, edit, update, and modify—whenever the appropriate user rights and permissions apply.

2. Why DMDG?

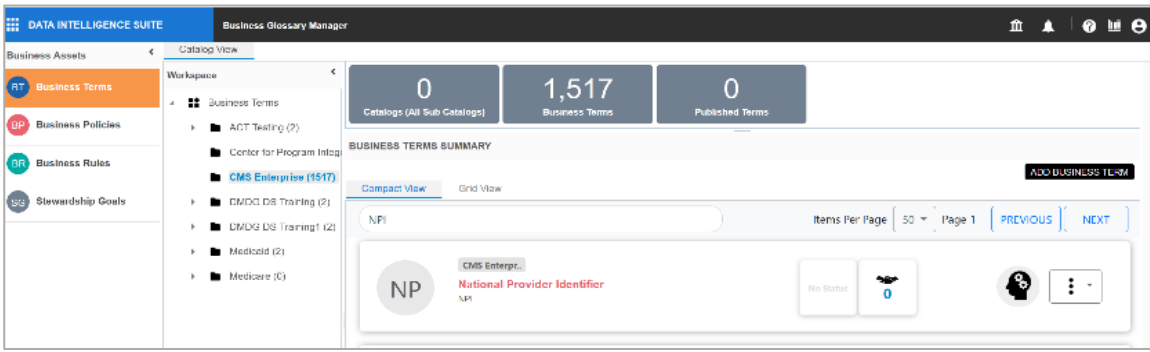
Why should The Center for Medicare and Medicaid Services (CMS) invest in a Metadata Management and Data Governance software application? In the Health Care industry, organizations are implementing Data Governance solutions to provide transparency into the physical structures of metadata that provide uniformity and transparency as metadata flows throughout organizations and Systems.



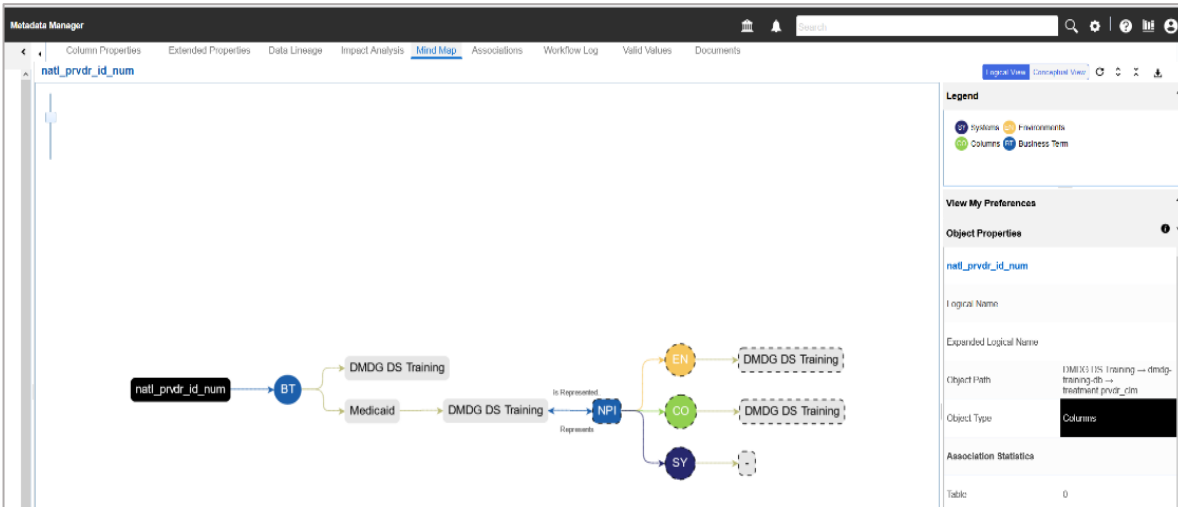
The need for a data strategy plan and Metadata Management arose from challenges of the National Provider Identifier (NPI) a 10-digit identifier representing a provide and shared with other health care providers for billing purposes. The Luhn Formula checksum algorithm is used to validate that the NPI number is correctly entered. NPI is maintained by the National Plan and Provider Enumeration System (NPPES) .See more details here <https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/downloads/NPI-What-You-Need-To-Know.pdf>

The NPI and its characteristics are well defined and the 2004 guidelines for “NPI” and “NPI Final Rule” can be downloaded here <https://www.cms.gov/Regulations-and-Guidance/Administrative-Simplification/NationalProvidentStand/Downloads/NPIcheckdigit.pdf>. But over the years, data stewards within the CMS Federal Contractor Systems, defined the data structures differently. When new technical integrations are added to a contractor System, the definitions need to be reworked to align with new technical requirements.

Defining Business Terms, Policies, and Rules within the Data Literacy Catalog of Erwin DI can solve the challenge of standardization and uniformity of metadata’s physical properties. The screenshot on the next page demonstrate how the CMS-Standard definition for “NPI” was added to the “CMS Enterprise” catalog of Business Terms within the Erwin DI Data Literacy module.



With this definition now built-in to the catalog, Application Data Stewards across all Systems can now create an “Association” between the Business Term, “NPI,” and technical assets such as environments, tables, or column. Once the Application Data Steward creates an “Association” between the column technical asset and the NPI business asset, the “Mind Map” functionality can generate an interactive diagram of all associated technical and business objects throughout the CMS Enterprise System.



This screenshot represents a Mind Map generated for all columns associated with “Provider NPI.”



See <https://www.cms.gov/research-statistics-data-Systems/DataArchitectureEngineering/CMSVocabulary> for more details about CMS-Standard Terms and the Name Checker Utility.

3. What is metadata?

Metadata is a data set that describes other data and gives information about it. The DMDG Erwin DI tool collects database metadata but does not look at the actual data stored in those structures. Physical database structures include the system, environment, schema, table, and column key values. For example, metadata within Erwin DI can be tagged and classified with Sensitive Data Indicators such as PHI or PII, but the actual data, let's say the provider claim number, or the beneficiary's social security number, is not stored in the System Catalog.

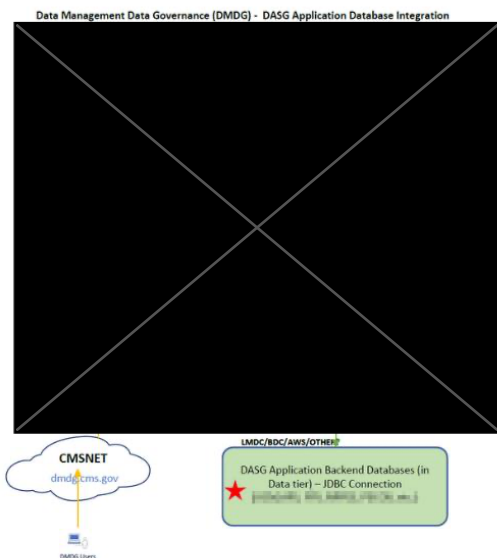
This is illustrated in the screenshot of the “Data Dictionary” results window on the next page. Data Dictionary captured everything about the “prvdr_clm” table from the original scan, including the table name, table definition, sensitive data flag, and table length.

#	Column Name	Logical Column Name	Column Comments	Column Definition	SDI Flag	Column Datatype	Length	Nullable Flag	Primary Key Flag	Column Storage Type	Column Alias	Column Class	Created By	Created Date	Modified
1	clm_id				🔒	bigint	8	N	Y				DL	2021-10-01 09:06:06.00	DL
2	clm_type_cd				🔒	char	10	Y	N				DL	2021-10-01 09:06:06.00	DL
3	clm_cd				🔒	char	10	Y	N				DL	2021-10-01 09:06:06.00	DL
4	clm_desc				🔒	varchar	128	Y	N				DL	2021-10-01 09:06:06.00	DL
5	clm_ansi				🔒	datetime	15	Y	N				DL	2021-10-01 09:06:06.00	DL
6	clm_dt				🔒	date	3	Y	N				DL	2021-10-01 09:06:06.00	DL
7	nat_prvdr_id_num				🔒	char	10	N	N				DL	2021-10-01 09:06:06.00	DL
8	bcch_bhon				🔒	char	15	N	N				DL	2021-10-01 09:06:06.00	DL
9	clm_type_d				🔒	smallint	2	N	N				DL	2021-10-01 09:06:06.00	DL

4. What is scanning?

Scanning metadata is known by several terms including “ingesting metadata,” “getting your metadata into the tool,” and “integration with DMDG.”

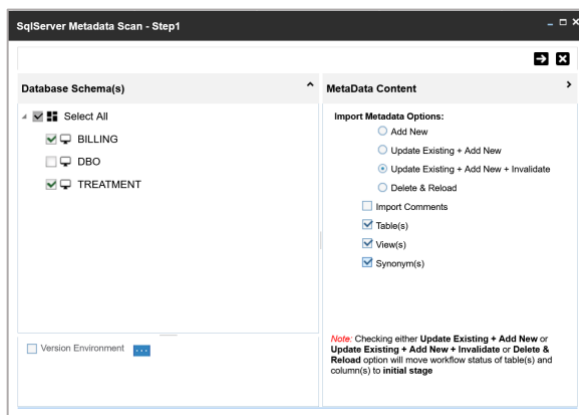
The diagram below illustrates a high-level integration of how each of the databases is ingested. The Erwin Data Intelligence “Standard Connector” is a JDBC-based connector that will initiate a connection to a System’s database. That connection requires Application Data Stewards to input the correct configuration parameters for their data tier into Metadata Manager. As a scan is performed, the Standard Connector will connect, query, and download the information maintained within that database’s metadata catalog. If an application System has more than one database, then there will be more than one Environment defined by the Metadata Manager for that System. The entire scanning process takes only a couple of minutes. The standard rule is: multiple databases result in multiple Environments within the System’s Erwin DI Data Catalog.



5. How to scan metadata

One of the common questions is “how long does it take?” The Erwin Data Intelligence Standard Connector queries and extracts the names and definitions of tables, columns, and views by directly connecting to the System’s database. The process takes only a few minutes.

From the Metadata Manager module of the Erwin DI tool, Application Data Stewards can configure the characteristics for what to include in the scan. Below is a screenshot of the **SqlServer Metadata Scan** dialog box.



DMDG recommends Application Data Stewards choose “Select All” from **Database Schema(s)** and the “Import Comments”, “Tables”, “Views” and “Synonyms” under **MetaData Content** “Import Metadata Options.” The rationale for selecting “Update Existing + Add New + Invalidate” is because *table deletes* in Erwin DI are “soft deletes”. Deleted tables can only be restored on the backend by the DMDG System Maintain team.

Also select “Table(s),” “View(s),” and “Synonym(s)” under **MetaData Content**, “Import Metadata Options.” Select this option *only if* comments exist within the physical data model. There is a known “Erwin Defect/Limitation” that results when comments are added manually after the initial scan and those comments are overwritten with blank values.

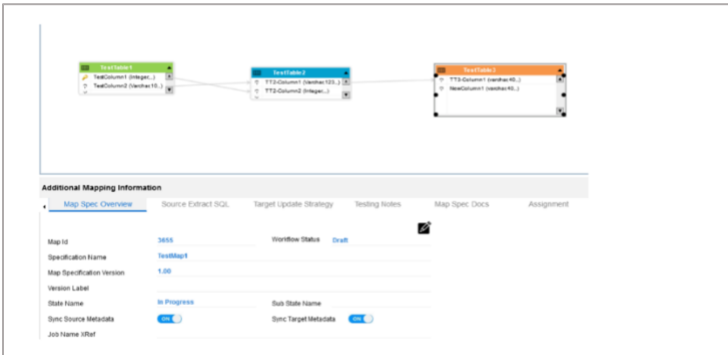
Application Data Stewards can also use "Version Environment", located at the bottom left of the **SQL MetaData Scan** dialog box. This option creates an archived record for the current metadata model within a scanned environment and allows changes to be observed over time.

If "Version Environment" is not selected for the first database scan, the Erwin DI product will not baseline this archive. Conversely, if changes to table or column objects are visible, the Erwin DI tool will not automatically version that environment. This behavior is directly controlled through the "Version Environment" selection. As applications promote changes using their change control processes, versioning should be used. In future versions of Erwin DI, the number of versions retained will be limited.

6. Mapping Specifications

When creating a mapping, always set "Sync Source Metadata" and "Sync Target Metadata" to "on." When a mapping specification is created between two columns, the Mapping Manager functionality creates its own copy of the Source and Target columns as well as their respective metadata values (type, length, precision). Without the Sync Source options "on", the metadata used in Mapping Manager is detached and will not receive Metadata scans operation unless the "Sync Source Metadata" and "Sync Target Metadata" are enabled.

If either of these flags is disabled, updates made to Source/Target columns in Metadata Manager will not be reflected in Mapping Specs. The "Sync" basically keeps Metadata Manager and Mapping Manager in alignment with column metadata. The screenshot below illustrates a "Map Spec Overview" (Mapping Specification Overview) where the "Sync Source Metadata" and "Sync Target Metadata" were enabled. These values can be easily updated after a map is created if accidentally set the wrong way.



In the complex mapping below, values cannot be updated because "Sync Source/Target Metadata" was disabled. Versioning in the Metadata Manager will not be synced. The Application Data Steward will need to create a new map.

