

DEPARTMENT OF HEALTH & HUMAN SERVICES  
Centers for Medicare & Medicaid Services  
7500 Security Boulevard, Mail Stop N2-14-26  
Baltimore, Maryland 21244-1850



## Central Data Administration

---

# Data Management Operating Procedures and Guidelines

---

May 3, 2011

*Prepared by:*  
*Office of Information Services/Enterprise Data Group (OIS/EDG)*  
*Division of Business Analysis & Operations (DBAO)*  
*7500 Security Boulevard, Baltimore, Maryland 21244-1850*

**Revision\Change Description History Log**

<b>Revision</b>	<b>Date</b>	<b>Revision/Change Description</b>	<b>Pages Affected</b>
Original		Baseline	N/A
Version 3.1	08/31/2009	Add cover page, history log, and table contents. Insert page numbers.	All
Version 4.0	11/30/2009	Add DM OP-031 and DM OP-032, revise cover page; remove DM G-021 and DM G-022; renumber DM G-023	pp 40 – 47 pp 70 – 88
Version 5.0	3/8/2010	Update DM OP-010 to include new operating procedure on foreign key definitions. Update DM OP-031 to address an additional UDP to capture National/International Standard Element Names Replace DM OP-023 with New Model Review Process	pp 16 pp 41 – 42 pp 29
Version 5.1	4/27/2010	Update DM OP-008 to further clarify entity definitions Update DM OP-012 to require representation terms to be consistent with definitions	p 14 p 19
Version 5.2	5/24/2010	Update DM G-010 and DM G-011 to differentiate between LDM/PDM vs. separate LDM and PDM	pp 55 - 57
Version 5.3	7/26/2010	Update DM OP-023 by adding URL for New Model Review Process	p 29
Version 5.4	9/27/2010	Remove DM OP-024 Add DM OP-033	p 30 p 45
Version 6.0	5/03/2011	Update DM OP-009 and DM OP-012 to reflect change in use of Object Class Terms	pp 15 – 16 pp 19 - 21

# Table of Contents

<b>DATA MANAGEMENT OPERATING PROCEDURES AND GUIDELINES .....</b>	<b>5</b>
<b>DATA MANAGEMENT OPERATING PROCEDURES .....</b>	<b>6</b>
DM OP-001 OPERATING PROCEDURE FOR REQUESTING DATA MANAGEMENT SERVICES .....	7
DM OP-002 OPERATING PROCEDURE FOR IDENTIFYING SYSTEM INTERFACES .....	8
DM OP-003: OPERATING PROCEDURE FOR DEVELOPING THE CONCEPTUAL DATA MODEL.....	9
DM OP-004: OPERATING PROCEDURE FOR ESTIMATING DATA MANAGEMENT SERVICE NEEDS .....	10
DM OP-005 OPERATING PROCEDURE FOR DEVELOPING THE LOGICAL DATA MODEL .....	11
DM OP-006 OPERATING PROCEDURE FOR REUSE OF ENTERPRISE ENTITIES, RELATIONSHIPS AND ATTRIBUTES .....	12
DM OP-007 OPERATING PROCEDURE FOR REUSE OF ENTERPRISE DATA RESOURCES.....	13
DM OP-008 OPERATING PROCEDURE FOR DEFINING DATA ENTITIES .....	14
DM OP-009 OPERATING PROCEDURE FOR NAMING DATA ENTITIES .....	15
DM OP-010 OPERATING PROCEDURE FOR DEFINING DATA ATTRIBUTES .....	16
DM OP-011 OPERATING PROCEDURE FOR ANALYZING TYPES OF DATA ATTRIBUTES .....	17
DM OP-012 OPERATING PROCEDURE FOR NAMING DATA ATTRIBUTES.....	18
DM OP-013 OPERATING PROCEDURE FOR MODELING DERIVED DATA .....	20
DM OP-015 OPERATING PROCEDURE FOR DEFINING RELATIONSHIPS, CARDINALITY, AND OPTIONALITY .....	21
DM OP-016 OPERATING PROCEDURE FOR ASSIGNING A PRIMARY IDENTIFIER.....	22
DM OP-017 OPERATING PROCEDURE FOR NORMALIZING THE PROJECT LOGICAL DATA MODEL .....	23
DM OP-018 OPERATING PROCEDURE FOR ASSIGNING DOMAIN VALUE RULES .....	24
DM OP-019 OPERATING PROCEDURE FOR CHECKING COMPLETENESS OF A PROJECT LOGICAL DATA MODEL .....	25
DM OP-020 OPERATING PROCEDURE FOR DOCUMENTING DATA ISSUES .....	26
DM OP-021 OPERATING PROCEDURE FOR ASSIGNING INFORMATION SECURITY CATEGORIES .....	27
DM OP-022 OPERATING PROCEDURE FOR GENERATING THE PROJECT METADATA REPOSITORY .....	28
DM OP-023 OPERATING PROCEDURE FOR PERFORMING A QUALITY REVIEW OF THE PROJECT LOGICAL DATA MODEL.....	29
DM OP-024 OPERATING PROCEDURE FOR APPROVAL OF DATA MODELS BASED ON BUSINESS DATA VIEWS .....	30
DM OP-025 OPERATING PROCEDURE FOR COLLECTING DATA STABILITY AND GROWTH INFORMATION .....	31
DM OP-026 OPERATING PROCEDURE FOR COMPLETING THE DATA MODELS .....	32
DM OP-027 OPERATING PROCEDURE FOR GRANTING PROJECT DATA MODEL LIBRARY ACCESS.....	33
DM OP-028 OPERATING PROCEDURE FOR NAMING AND DEFINING DATA MODELS.....	34
DM OP-029 OPERATING PROCEDURE FOR SELECTING MODEL TYPE .....	35
DM OP-030 OPERATING PROCEDURE FOR CONDUCTING A PHYSICAL DATA MODEL REVIEW.....	36
DM OP-031: OPERATING PROCEDURE FOR CAPTURING THE STANDARD LOGICAL DATA MODEL METADATA .....	37
DM OP-032: OPERATING PROCEDURE FOR CAPTURING THE STANDARD PHYSICAL DATA MODEL METADATA.....	41
DM OP-033: OPERATING PROCEDURE FOR RECORDING DATA MODEL CHANGES .....	45
<b>DATA MANAGEMENT GUIDELINES .....</b>	<b>47</b>
DM G-001 GUIDELINE FOR ASSIGNING DATA ANALYSTS.....	48
DM G-002 GUIDELINE FOR CONDUCTING THE LOGICAL DATA DESIGN KICKOFF MEETING .....	49
DM G-003 GUIDELINE FOR MONITORING SCHEDULE AND COST OF DATA SERVICES .....	50
DM G-004 GUIDELINE FOR DESIGNATING REPRESENTATION TERM AND DATA TYPE .....	51
DM G-005 GUIDELINE FOR CONSTRUCTING PHYSICAL DATA NAMES .....	52
DM G-006 STANDARD FOR ASSIGNING DATE FORMATS.....	53
DM G-007 GUIDELINE FOR CONDUCTING THE DATA ARCHITECTURE REVIEW .....	54
DM G-008 GUIDELINE FOR MAPPING DATA REQUIREMENT SOURCES.....	55
DM G-009 GUIDELINE FOR PREPARING THE PROJECT LOGICAL DATA MODEL FOR PHYSICAL DESIGN .....	56
DM G-010 GUIDELINE FOR CONSTRUCTING PHYSICAL TABLE OR FILES NAMES .....	57
DM G-011 GUIDELINE FOR CONSTRUCTING PHYSICAL COLUMN OR ELEMENT NAMES .....	58
DM G-012 GUIDELINES FOR ASSIGNING DATA PLANNING PARTICIPANTS.....	60
DM G-013 GUIDELINES FOR CONDUCTING ENTERPRISE DATA WORK SESSIONS .....	61
DM G-014 GUIDELINE FOR CREATING A BUSINESS PROCESS DIAGRAM.....	62
DM G-015 GUIDELINES FOR DEFINING A SUBJECT AREA .....	63
DM G-016A GUIDELINES FOR ANALYZING AND APPROVING STANDARD TERMS.....	64
DM G-016B GUIDELINES FOR ABBREVIATING A STANDARD TERM.....	66

## Data Management Operating Procedures and Guidelines

DM G-017 GUIDELINES FOR REFERENCE TABLES .....	67
DM G-018 GUIDELINES FOR MANAGING DATA VALUES THROUGH PHYSICAL CONSTRAINTS .....	68
DM G-019 GUIDELINE FOR MODELING SUPERTYPES AND SUBTYPES .....	69
DM G-020 GUIDELINE FOR USING ERWIN COMPLETE COMPARE TO CREATE A BRAND NEW DATA MODEL .....	70
DM G-021 GUIDELINE FOR NAMING REFERENCE ENTITIES .....	71

## **Data Management Operating Procedures and Guidelines**

The full text of operating procedures, guidelines, and standards, including their underlying *rationale*, and *level of enforcement* are enclosed for your reference.

Numbering schemes follow this format with *nnn* as a sequence number (e.g., 001):

DM OP *nnn* (Operating Procedure)

Operating procedures mandate certain actions that are enforceable for specified IT Project.

DM G *nnn* (guideline)

Guidelines provide best practice methods, and are optional.

## **Data Management Operating Procedures**

## ***DM OP-001 Operating Procedure for Requesting Data Management Services***

### ***Rationale:***

The Data Services Manager needs certain information to adequately plan and assign data management service resources in support of CMS projects. That information is collected through the *Data Management Service Request Form*, which is designed to capture it in a clear and organized format.

### ***Operating Procedure:***

A request for data design services shall be initiated through submission of the *Data Management Service Request Form*.

## ***DM OP-002 Operating Procedure for Identifying System Interfaces***

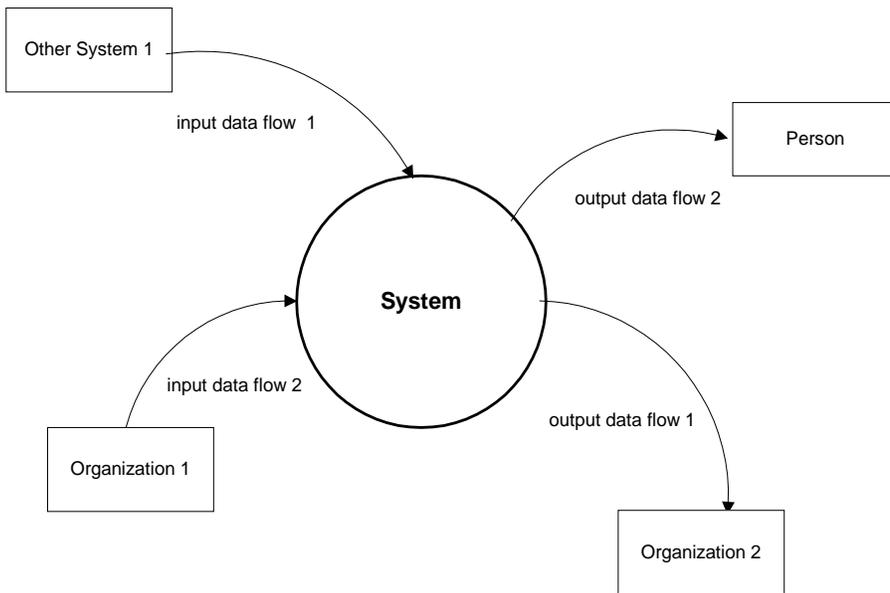
### ***Rationale:***

The System Context Diagram concisely illustrates the system boundary in relation to its environment, and documents the actual or planned data flows into and out of the system.

### ***Operating Procedure:***

Business application interfaces shall be identified using a System Context Diagram. See example.

### **System Context Diagram**



### ***DM OP-003: Operating Procedure for Developing the Conceptual Data Model***

***Rationale:***

The Conceptual Data Model diagram illustrates the major entities about which the business enterprise needs information. This diagram will assist in the identification of existing data sources and help in determining the need for new entities.

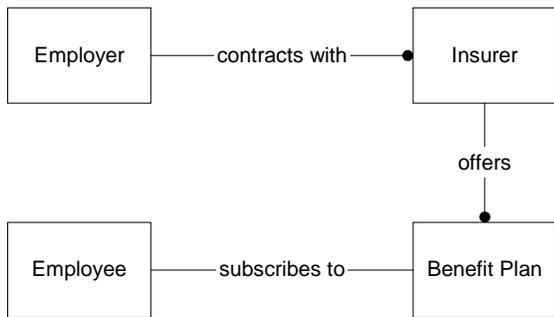
***Operating Procedure:***

Business applications that create, update, or replicate data shall have high-level data requirements documented in a Conceptual Data Model.

Conceptual Data Models shall represent business entity relationships that are within the scope of the target business function(s) in a manner depicted by the *Example Conceptual Data Model*.

Major IT projects may require Data Architect review

#### ***Example Conceptual Data Model***



**Conceptual Data Model Features:**

- Includes the important entities and the relationships among them.
- No attribute is specified.
- No primary key is specified

**Business Rule:**

An employer contracts with one or more insurance companies to offer health care benefits to any employee who elects to subscribe.

***DM OP-004: Operating Procedure for Estimating Data Management Service Needs***

*Rationale:*

The project sizing and estimated cost provides the business enterprise the information for calculation of potential return on investment (*ROI*).

*Operating Procedure:*

Prior to starting data management services, the Project Sponsor shall ensure development of an estimate of the effort (e.g., dollar amount/hours, data management resources) and schedule of the data services work necessary to satisfy project needs. Data Management services shall be negotiated with the Data Services Manager.

## ***DM OP-005 Operating Procedure for Developing the Logical Data Model***

### *Rationale:*

Business requirements are to be documented in a standard notation for consistent documentation with a view for prospective data reuse, and effective communication of business data requirements.

### *Operating Procedure:*

Business applications that create, update, or replicate data shall have their data requirements documented in a Logical Data Model. The Logical Data Model shall be developed in the standard data modeling tool, or in a format that can be directly imported into the standard data modeling tool.

## ***DM OP-006 Operating Procedure for Reuse of Enterprise Entities, Relationships and Attributes***

### *Rationale:*

Shared business data artifacts must be managed according to enterprise data stewardship objectives.

### *Operating Procedure:*

Project Data Analysts shall use shared model objects where available rather than duplicating them for exclusive use by each project. Reuse of enterprise entities and attributes requires observance of the following requirements:

1. New data definitions may not redundantly define information already in the Enterprise Metadata Repository. The goal is to ensure that a new data artifact does not duplicate an existing one.
2. When a project selects entity or attributes from the Enterprise Data Model, a log of changes must be kept throughout the project lifecycle. The log must include entries for the following instances:
  - a. Altered entity and attribute definitions
  - b. Changes to cardinality or optionality
  - c. Deletions, migrations, or changes of any kind

## ***DM OP-007 Operating Procedure for Reuse of Enterprise Data Resources***

### *Rationale:*

Shared data resources must be managed according to enterprise data management objectives.

### *Operating Procedure:*

Project Data Analysts shall make use of existing enterprise data resources. Reuse of enterprise data resources requires the development of a project Data Source Plan, which identifies candidate sources for the data required to satisfy the project's business requirements. For each candidate data source, the following shall be documented:

1. Name
2. Description
3. Business Owner or steward
4. Project data needs satisfied
5. Limitations (e.g., data quality issues)
6. Constraints (e.g., operational impact of using source)

## *DM OP-008 Operating Procedure for Defining Data Entities*

### *Rationale:*

A well-formed definition makes a data object's contents apparent to business users.

### *Operating Procedure:*

1. An entity shall be defined in a manner that distinguishes its unique role within the business enterprise.
2. The definition shall be clear, concise, and unambiguous. Examples or exclusions may be added to the definition to improve clarity.
3. The definition shall describe a single occurrence of the entity.
  - a. The definition shall start with either 'A' or 'An', e.g. "A person who...", or "An individual that..."
  - b. The definition shall not start with a verb, e.g. "Tracks the audits...", "Captures beneficiary...", etc.
4. Since data entities are not generally technology concepts, the definition may not include references to application privileges, interface behavior or internal database auditing.
5. The definition shall not include references to technology or media. For example references to "tapes" or "disks" are not appropriate in the definition of a business entity.
6. The definition shall exclude references to time-bound or process-bound circumstances. For example, "the information the agency receives quarterly" is not included in a definition.
7. The entity definition may not include acronyms or abbreviations not found in the *Standard and Abbreviations Terms List*.
8. Subject matter experts who understand the business meaning, in conjunction with the data analyst, should develop definitions. If the data analyst develops the definition, it is important that subject matter experts review and confirm the accuracy of the definition.

## *DM OP-009 Operating Procedure for Naming Data Entities*

### *Rationale:*

A good data entity name is meaningful and self-documenting. It provides the first source of information about the entity's purpose and contents.

### *Operating Procedure:*

1. Entity names shall be unique throughout the Enterprise Data Model and any individual Project Data Model. This requires that a new entity name not duplicate existing entity names in the Enterprise Data Model or the individual Project Data Model.
2. An entity name shall be stated in singular form.
3. Base the entity name on the respective entity definition.
4. An entity name may contain alphabetic characters A-Z and spaces. It may not contain numbers or special characters, including the slash (/) and hyphen (-).
5. An entity name is composed of the following types of terms:

<b>Term Type</b>	<b>Cardinality</b>
Object Class	1
Qualifier	0 - n

There is no required ordering of terms. The list of valid terms can be found on the CMS Data Administration web site:

<http://www.cms.gov/DataAdmin/>

- **Object Class Term** is selected to categorize the entity into a higher level group, e.g. 'Beneficiary', 'Claim', 'Provider'
  - **Qualifier Terms** are included to provide more meaning, differentiating between entities with the same Object Class Term.
6. Format entity names in mixed case (the first letter of each term is in uppercase, the remaining letters in the term are in lowercase) throughout the model
  7. Avoid using excessively long business names for entities. If a standard business name exceeds the maximum length of 120 characters (including embedded spaces), progressively remove the least significant *qualifier term(s)*.
  8. Use a *space* as the delimiter between terms.
  9. Use of acronyms in entity names is prohibited in the following cases:
    - (1) Enterprise Data Models
    - (2) As an Object Class Term in any logical data model.
  10. All approved acronyms are found in the [CMS Standard Terms and Abbreviation list](#).
  11. Abbreviations that are not acronyms, shall only be used when it is necessary to fit the entity name within the 120 character limitation.

## *DM OP-010 Operating Procedure for Defining Data Attributes*

### *Rationale:*

A well-formed definition makes an attribute's contents apparent to business users.

### *Operating Procedure:*

1. An attribute shall be defined in a manner that distinguishes its unique role within the business enterprise.
2. An attribute definition shall be clear, concise, and unambiguous. Examples or exclusions may be added to the definition to improve clarity.
3. The definition shall describe a single occurrence of the attribute.
4. The definition shall not include references to technology or media. For example references to "tapes" or "disks" are not appropriate in the definition of a business attribute.
5. The definition shall exclude references to time-bound or process-bound circumstances. For example, "the information the agency receives quarterly" is not included in a definition.
6. The attribute definition may include well-recognized agency acronyms but no other acronyms or abbreviations.
7. Definitions should be developed by subject matter experts who understand the business meaning, in conjunction with the data analyst. If the data analyst develops the definition, it is important that subject matter experts review and confirm the accuracy of the definition.
8. When a relationship is created, a foreign key attribute automatically inherits its definition from the parent attribute. Do not modify this definition to explain why the foreign key exists in the child entity. Create a definition for the relationship.
9. An attribute definition should:
  - a. State the essential meaning of the concept
  - b. Be precise and unambiguous
  - c. Be concise
  - d. Be able to stand alone
  - e. Be expressed without embedding rationale, functional usage, domain information, or procedural information
  - f. Avoid circular reasoning
  - g. Use the same terminology and consistent structure for related definitions

## ***DM OP-011 Operating Procedure for Analyzing Types of Data Attributes***

### *Rationale:*

Attributes can be named and defined with additional consideration for how they are used and therefore, enhance their recognition and capability for storing business data values.

### *Operating Procedure:*

Define the attribute in the Logical Data model by the following criteria:

1. **Prime** attributes are basic business facts. They are always added to the model.
2. **Derived** attributes can be formulated using values from other attributes. They may be added to the Project Logical Data Model if useful to business communication but not necessarily transformed to physical data design. See *220-15 Modeling Derived Attributes* for additional information about inclusion of derived data in the Project Logical Data Model.
3. **Transaction / Interface transaction elements** are included only if they are to be transformed to the physical data design.
4. **Physical control data**, such as “next record number”, are not included in the logical data model.

## DM OP-012 Operating Procedure for Naming Data Attributes

### Rationale:

A good data attribute name is meaningful and self-documenting. It provides the first source of information about the data attribute's purpose and contents.

### Operating Procedure:

1. A data attribute's name shall be based on its definition. The name suggests the business meaning of the attribute and reflects the most important words in its definition. Match these important words with the terms in the [CMS Standard Terms and Abbreviation list](#).
2. A data attribute name consists of the following types of terms:

Term Type	Cardinality
Property	0 - n
Qualifier	0 - n
Representation	1

- **Property Term** – Identifies the property that the attribute represents, e.g. 'Color'. Often the Representation Term also serves as the Property Term, e.g. 'Name', so there may not be a separate Property Term. The Property Term must be spelled out, not abbreviated. The use of approved acronyms is allowed.
  - **Qualifier Term** – Further describes the attribute, allowing differentiation between attributes with the same Property and Representation Terms, e.g. in 'First Name' and 'Last Name' 'First' and 'Last' are Qualifier Terms. The Qualifier Term must be spelled out, not abbreviated. The use of approved acronyms is allowed.
  - **Representation Term** - Identifies the type of value that the attribute takes on, e.g. 'Code', 'Name', 'Number'. The Representation Term is mandatory, and must be the last term in the attribute name. The Representation Term must be spelled out, not abbreviated. Do not use an acronym at the end of an attribute name even if the last term of the acronym is an approved Representation Term. The Representation Term should be consistent with the definition of the attribute. For example, if the definition of the attribute is "the date the Employee is born", then the Representation Term should be 'Date'.
3. In addition to the terms described above, an attribute that is part of the primary key of an entity must start with the entity name. The purpose of this is so that it will be unique when it appears as foreign key in another entity because of a relationship. This only applies when the attribute appears for the first time in the entity, i.e. is not a foreign key inherited from another entity. For example, the key of Beneficiary should be 'Beneficiary Identifier', not just 'Identifier'. For entities that end with the term 'Reference', the primary key attribute name can omit it, for example the primary key for the entity 'Status Reference' should be 'Status Code', not 'Status Reference Code'.
  4. Format attribute names in mixed case (the first letter of each term is in upper case, the remaining letters are in lower case) throughout the model
  5. An attribute name shall be stated in singular form.

6. An attribute name must contain only alphabetic characters, upper and lower case.
7. An attribute name must not exceed 120 characters. If a standard business name exceeds 120 characters, progressively remove the least significant Property Term or Qualifier Term.
8. Use of acronyms in attribute names is prohibited in the following cases:
  - (a) Enterprise Data Models
  - (b) As a Representation Term in any logical data model.
9. All approved acronyms are found in the [CMS Standard Terms and Abbreviation](#) list.

## *DM OP-013 Operating Procedure for Modeling Derived Data*

### *Rationale:*

An example of a useful derived data element is the balance of a bank account. You could derive it from all previous credit and debit statements, since the start balance is 0. However, it would be very unpractical to do this every time the account balance is needed.

Another example of derived data is *totals* and *statistics*, for example the number of *Beneficiaries* in a state.

In the above examples, the data might be stored as part of the record to save processing time. However, one must be careful with storing derived data. When something changes in the source data, the derived data must change accordingly.

If the source data is very dynamic, it's better to just derive it every time.

### *Operating Procedure:*

1. Derived data may be included in the Logical Data Model as a convenient way to communicate the meaning of calculated business data.
2. Derived data may or may not be included in the Physical Data Model, depending on the benefit to actually store the derived product.
3. Use the following criteria in guiding whether to store derived data.

Pro → Store Derived Attribute	Con → Derive Attribute at processing time
Original factors not always available	Original Factors always available
Retrieval performance is important	Updates are frequent
Original factors are static	Original factors are dynamic
Up-to-date accuracy not important	Up-to-date accuracy essential
Derivation is costly or difficult	Derivation is cheap and easy

4. If derived information is included as an attribute, it must be noted as “derived” in the attribute definition and the derivation formula must be included in the “Attribute Derivation Text” UDP for the derived attribute in the logical data model. Reference Central Data Administration Standard Data Modeling Tool Use document.

## ***DM OP-015 Operating Procedure for Defining Relationships, Cardinality, and Optionality***

### *Rationale:*

The labeling of entity relationships in the data model aids the understanding of model details. Also, cardinality labels define the minimum and maximum number of entity occurrences that can participate in a relationship.

### *Operating Procedure:*

1. Redundant relationships between two entities are not allowed.
2. A relationship has only one possible state.
3. Each relationship must be named with a verb or verb phrase to express the relationship in the parent-child direction.
4. Relationship names may be expressed for each direction of the relationship.
5. Non-descriptive relationship names such as “belongs to”, “may have”, “is associated with” or “is a part of” should not be used.
6. Cardinality must be specified for each relationship.
7. Document the description of the entity parent to child relationship to clarify the relationship and describe the conditions of the relationship and any special conditions under which it exists.

## ***DM OP-016 Operating Procedure for Assigning a Primary Identifier***

### *Rationale:*

An entity must be defined in a way that the business can distinguish unique occurrences.

### *Operating Procedure:*

1. In the Project Logical Model, every entity type should have a unique identifier composed as one of the following options:
  - a. One attribute
  - b. Multiple attributes
  - c. An identifying relationship and one attribute
  - d. Two identifying relationships
2. An assigned number may be used as an identifier. This is rarely justified but is permitted for practicality when it represents a unique abbreviated way of referring to an occurrence of an entity type or needed to compensate for other unusual circumstances.
3. Examples:
  - a. A generated sequential number may be used to identify a particular entity, where the combination of the identifying attributes and relationship is much too cumbersome for practical use as an identifier.
  - b. During a company merger, records from two sources are to be combined and their existing primary identifiers might result in duplicate identifiers without an assigned number, which acts as a tiebreaker.

## ***DM OP-017 Operating Procedure for Normalizing the Project Logical Data Model***

### *Rationale:*

The purpose of data normalization is to ensure that (1) each attribute is assigned to the appropriate entity (2) redundant storage of information is minimized and (3) data structures are stable and receptive to updates, reducing the risk for introduction of anomalies.

### *Operating Procedure:*

1. The Project Logical Model shall be normalized to *First Normal Form*. All attributes of every entity must be atomic, indivisible, and not part of a repeating group.
2. The Project Logical Model shall be normalized to *Second Normal Form*. All attributes of every entity must depend on the full primary key of that entity only.
3. The Project Logical Model shall be normalized to *Third Normal Form (3NF)*. All attributes of every entity must depend on the full primary identifier of that entity only and not be functionally dependent on some other attribute.
4. The Project Logical Model may be normalized to *Fourth Normal Form*. The model is in *3NF* and has no sets of multi-valued dependencies.
5. The Project Logical Model may be normalized to *Fifth Normal Form* such that every join dependency is made through candidate identifiers.

***DM OP-018 Operating Procedure for Assigning Domain Value Rules***

*Rationale:*

Value domain assignments support business data definition and quality by restricting data to a certain set of permitted values.

*Operating Procedure:*

When the business operation depends on data to be within a discrete set of permissible values, the permitted attribute value range must be documented and managed.

## ***DM OP-019 Operating Procedure for Checking Completeness of a Project Logical Data Model***

### *Rationale:*

A Project Logical Data Model that conforms to all relevant standards and operating procedures supports the agency's long-term data architecture goals for creation of stable, reusable data artifacts.

### *Operating Procedure:*

Verify that:

1. The entities in the logical model represent all the things of interest within the scope of the business mode.
2. The logical model supports every data need identified in the functional business requirements.
3. The entities and attributes in the logical model have well-formed business definitions.
4. The model adheres to naming standards, notation, and security setting requirements.
5. The logical model is free of data redundancies, incorrect keys, and relationship errors.
6. The logical model is free of derived data or that such data is clearly marked as derived to ensure conscious transference to the physical design layer.
7. The logical model component names are consistent with standard naming conventions.
8. The logical model in is Third Normal Form.
9. The logical model is independent of physical design constraints.
10. Text that identifies the model diagram is located in the upper left corner: the application name; subject area name, if applicable; update date; and modeler identification.

## *DM OP-020 Operating Procedure for Documenting Data Issues*

### *Rationale:*

Sometimes it is not possible to express a data issue in the data modeling tool however, it is still necessary to communicate the issue and track its resolution.

### *Operating Procedure:*

Use the *Data Issues Notes* to report the following:

1. Document business rules that cannot be expressed using the data modeling tool.
2. Document data quality issues of any existing project data that might be in opposition to stated business rules.
3. Document unusual database physical platform location requirements.
4. Document data conversion challenges and any other issues that require special attention from DBA or business users.

Document Data Issues following the example below, including the underlined sections:

### Example

*Data Issue:* The current production table column *BENEFICARY. INDV\_WGHT\_AMT* has a defined data type of CHAR(3), which has resulted in entry and storage of invalid weight data. Project “*CMIS: Contractor Management Information System*” requires that the column be changed to *smallint*. Therefore existing values that are non compliant with the revision, must be corrected before it can be reloaded to the revised table column.

*Date Recorded:* 2004/08/24

*Resolution:* A conversion program has been developed to correct invalid weight data in table column *BENEFICARY. INDV\_WGHT\_AMT*. Doug Nichols, the project manager will coordinate correction of invalid data with the assigned database administrator, prior to production change implementation.

*Date Resolved:* 2004/08/25 by *Mable Hanks, assigned Project Data Analyst.*

***DM OP-021 Operating Procedure for Assigning Information Security Categories***

*Rationale:*

The Federal Information Security Act (FISMA) requires all federal agencies to document the appropriate security categories of asset and operational data.

*Operating Procedure:*

1. Enterprise data shall be assigned with appropriate security settings in compliance with *FIPS Publication 199 – Standards for Security Categorization of Federal Information and Information Systems*.
2. Each entity in a Logical Data Model shall be assigned a potential impact level for the following security objectives: confidentiality, integrity, and availability.
3. Further guidance is included in the following table.

<b>Security Objective</b>	<b>LOW</b>	<b>MODERATE</b>	<b>HIGH</b>
<p><b><i>Confidentiality</i></b> Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. [44 USC,SEC.3542]</p>	<p>The authorized disclosure of information could be expected to have a <b>limited</b> adverse effect on organizational operations, organizational assets, or individuals.</p>	<p>The authorized disclosure of information could be expected to have a <b>serious</b> adverse effect on organizational operations, organizational assets, or individuals</p>	<p>The authorized disclosure of information could be expected to have a <b>severe or catastrophic</b> adverse effect on organizational operations, organizational assets, or individuals</p>
<p><b><i>Integrity</i></b> Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity. [44 USC,SEC.3542]</p>	<p>The unauthorized modification or destruction of information could be expected to have a <b>limited</b> adverse effect on organizational operations, organization assets, or individuals.</p>	<p>The unauthorized modification or destruction of information could be expected to have a <b>serious</b> adverse effect on organizational operations, organization assets, or individuals.</p>	<p>The unauthorized modification or destruction of information could be expected to have a <b>severe or catastrophic</b> adverse effect on organizational operations, organization assets, or individuals.</p>
<p><b><i>Availability</i></b> Ensuring timely and reliable access to and use of information. [44 USC,SEC.3542]</p>	<p>The disruption of access to or use of information or an information system could be expected to have a <b>limited</b> adverse effect on organizational operations, organizational assets, or individuals.</p>	<p>The disruption of access to or use of information or an information system could be expected to have a <b>serious</b> adverse effect on organizational operations, organizational assets, or individuals.</p>	<p>The disruption of access to or use of information or an information system could be expected to have a <b>severe or catastrophic</b> adverse effect on organizational operations, organizational assets, or individuals.</p>

## ***DM OP-022 Operating Procedure for Generating the Project Metadata Repository***

### *Rationale:*

The Metadata Repository documents the understanding of business data and facilitates the communication about the information that is represented in the Logical Data Model.

### *Operating Procedure:*

The Metadata Repository holds the following information:

- 1. For each entity**
  - a. Entity Name
  - b. Entity Definition
  - c. Degree of required Confidentiality
  - d. Degree of required Integrity
  - e. Degree of required Availability
  
- 2. For each attribute**
  - a. Attribute Name
  - b. Definition
  - c. Primary Key Indicator
  - d. Foreign Key Indicator
  - e. Data type
  - f. Valid Values
  - g. Edit Criteria
  - h. Business transformation rules

***DM OP-023 Operating Procedure for Performing a Quality Review of the Project Logical Data Model***

*This Operating Procedure has been replaced by the Model Review Process. Please see the Logical Design page [http://www.cms.gov/DataAdmin/03\\_LogicalDataDesign.asp#TopOfPage](http://www.cms.gov/DataAdmin/03_LogicalDataDesign.asp#TopOfPage).*

***DM OP-024 Operating Procedure for Approval of Data Models based on Business Data Views***

*This Operating Procedure has been replaced by the Model Review Process. Please see the Logical Design page [http://www.cms.gov/DataAdmin/03\\_LogicalDataDesign.asp#TopOfPage](http://www.cms.gov/DataAdmin/03_LogicalDataDesign.asp#TopOfPage).*

## ***DM OP-025 Operating Procedure for Collecting Data Stability and Growth Information***

### *Rationale:*

Based on predictions by subject matter experts (SMEs), resources can be better chosen to accommodate near-term volume expectations. This procedure encourages collection of data stability and growth information that might be used by the Data Base Manager in selection of the appropriate Data Base Management System (DBMS) platform. (This information serves only as preliminary prediction. Additional, detailed planning for physical storage allocations will be performed by the assigned data base administrator (DBA) in the course of physical database implementation.)

### *Operating Procedure:*

This information is typically captured on the ERWIN Volumetric Sizing Information table property tab within the PDM (Physical Data Model) for each table. Document the expected growth rate of a new table predicted by the SMEs, as an annual increase in the number of rows. For example if the SME predicts 100,000 more records in 2 years, the estimated annual growth is 50,000 records which can be specified in the Grow By property value in the Erwin Volumetric Sizing Information table property tab.

***DM OP-026 Operating Procedure for Completing the Data Models***

*Rationale:*

The data models are stored and cataloged to benefit ongoing enterprise data analysis and business data analysis.

*Operating Procedure:*

All data models shall be stored and cataloged in DDMSS model management libraries, according to their state of completion.

All models shall be appropriately stored when work is completed (or halted in an incomplete or unapproved status).

## *DM OP-027 Operating Procedure for Granting Project Data Model Library Access*

### *Rationale:*

Storing models in DDMSS libraries provides protected storage for data models and data management documents.

### *Operating Procedure:*

1. Each project shall have designated model management libraries:
  - a. 2 libraries in Model Manager; one for development models, the other for production models.
  - b. 1 folder in the DDMSS Change Control library to hold project Data Issue Documents and Change Control Documents.
2. Full update privileges for development libraries and viewing privileges to project production libraries are available to authorized Project Data Analysts and Data Base Administrators.
3. Update privileges for Model Manager and DDMSS production model libraries are exclusive to the Model Manager Administrator.
4. Assign a name and provide a description for each Model Manager library in the following manner:
5. The library name consists of: *application acronym + application business name + phase designation + approval status*.
  - a. Example: *CMIS: Contractor Management Information System Approved*.
6. The application library description should be brief and may include the names of the assigned data analysts, project manager, and government task leaders (GTL).
  - a. Example: *As an intended replacement for the CROWD system, CMIS allows CMS to effectively manage, monitor, and report on the performance of its Medicare fee-for-service contractors. Project Manager is Doug Nichols of Acme Systems, Inc. The Project Data Analyst is Mable Hanks. The Project Database Administrator is Nancy Owens.*
7. Create a folder in the *DDMSS Change Control Library*.
8. The folder name convention is: *application acronym + application business name + phase designation + YYYYMMDD*.
  - a. Example: *CMIS: Contractor Management Information System 20030401*.

## ***DM OP-028 Operating Procedure for Naming and Defining Data Models***

### *Rationale:*

Consistent names and description formats are important to the cataloging of data models, managing them as data architecture resources, and exploiting their potential for reuse.

### *Operating Procedure:*

A data model is to be named in the following manner:

*system acronym + (“relational”/”dimensional”)+ model type (EDM/CDM/LDM/PDM) + approval date (or the storage date for models in development) in yyymmdd format.*

Example: **CMIS relational LDM 20040726.**

The model definition describes the purpose and status of the model in a few sentences of text. It is optional.

Example: **Current approved LDM.**

## DM OP-029 Operating Procedure for Selecting Model Type

### Rationale:

The CMS standard data modeling tool, *All-Fusion ERwin Data Modeler*, provides convenient functions that help the data modeler to create models that support the various levels of CMS data architecture.

### Operating Procedure:

Select the model type according to project needs. Logical models capture data requirements. Physical models support the development of a database design in parallel with or separately from the associated conceptual or logical model development.

1. The following table lists the allowed model source relationships for logical and physical models:

Parent Model Type	Child Model Type	Notes
Logical	Logical	For example, enterprise to project logical, or project conceptual to project logical
Logical	Physical	You can derive a physical model from a logical model. The result is a pair of linked models with the logical model as the parent and the physical model as the child
Physical	Physical	For example, generic physical to target server specific physical

2. A third model type, logical/physical, created by ERwin version 3.5.2 or earlier, combines one logical model and one physical model that are tightly coupled. Use a logical/physical model when you require only one physical implementation for a given logical model and do not need to continue logical modeling while the physical design is being implemented. Logical/physical models can also be created by reverse engineering an implemented database or a DDL script. Logical/physical models are used in design layer architectures, but their usage is a little different. The allowed model source relationships that result from splitting a logical/physical model are:

Parent Model Type	Child Model Type	Notes
Logical/Physical	Logical	The logical level and model properties such as the naming standards model and data type standards model may be passed to the child model
Logical/Physical	Physical	The physical level, and model properties such as the naming standards model and data standards model may be passed to the child model

***DM OP-030 Operating Procedure for Conducting a Physical Data Model Review***

*This Operating Procedure is no longer used.*

## ***DM OP-031: Operating Procedure for Capturing the Standard Logical Data Model Metadata***

### *Rationale:*

Capturing a standard set of metadata associated with the Logical Data Model provides the necessary details to foster communication within and across projects.

### *Operating Procedure:*

The following metadata shall be captured for each Model, Entity and Attribute in the LDM.

### **Model Metadata Standards**

The following identifies and describes the required CMS Standard Model-Level Metadata that shall be captured for each Logical Data Model.

<b>CMS Standard Model-Level Metadata</b>	<b>Model-Level Metadata Description</b>
LDM DA Signoff Date	The date when this version of the logical data model received its Central DA sign-off.
Model Business Contact Name	The name of the organization/ person who is responsible for approving the definitions in the model. <i>Example: R. R Kirk (CMS/CMM)</i>
Model Central DA Name	The name of the person who is responsible for certifying the compliance of the model components according to CMS Data Administration Standards.
Model Create Date	The date recorded by Local DA for the original Model.
Model Name	See: <a href="#">DM OP- 028 Operating Procedure for Naming and Defining Data Model</a>
Model Author	The first and last name of the Local or Central Data Architect as well as the company name of the person responsible for creating and/or maintaining the data model. <i>Example: <b>Bob Smith, XYZ Company</b></i>
Model Type	See: <a href="#">DM OP-029 Operating Procedure for Selecting Model Type.</a>
Model Definition	See: <a href="#">DM OP- 028 Operating Procedure for Naming and Defining Data Model</a>
LDM Model Last change Description	A brief narrative that summarizes the nature of the changes resulting in the current version of the logical data model.
Model Central DBA Name	The name of the person who is responsible for certifying the compliance of database components according to the CMS Database Administration Standards.

**Entity Metadata Standards**

The following identifies and describes the required CMS Standard Entity -Level Metadata that shall be captured for each entity identified in the Logical Data Model.

CMS Standard Entity-Level Metadata	Entity-Level Metadata Description
Entity Name	See: <a href="#">DM OP-009 Operating Procedure for Naming Data Entities</a>
Entity Definition	See: <a href="#">DM OP-008 Operating Procedure for Defining Data Entities</a>
Entity Requirement ID	<p>A reference to the requirement(s) or change request identifier(s) that justify the existence of the entity in the model.</p> <p>Valid entry formats are as follows:</p> <ul style="list-style-type: none"> <li>(1) DOORS or requirement document Tracking ID, use – BR-#### for business requirement FR-##### for functional requirement, SR-#### for system requirements.</li> <li>(2) Change requests - use CR-####.</li> <li>(3) Data change requests - use DR-#####.</li> <li>(4) Remedy Tickets - use RT_#####.</li> <li>(5) MAPD - use MAPD_####.</li> </ul>
Entity Security Category Description	See: <a href="#">DM OP-021 Operating Procedure for Assigning Information Security Categories</a>
Logical Only Entity Indicator	Required for all logical only entities
Entity Business Contact Name	<p>Required if the contact is different from the Model Business Contact Name. The name of the organization/ person who is responsible for approving the definitions of the Entity and its contained Attributes.</p> <p><i>Example: R. R Kirk (CMS/CMM)</i></p>
Entity CDA Standard Name	The CDA-approved name of the entity, if different from the entity’s name in the model. This standard logical name must be used when creating a new data entity to represent the identical concept in a new or existing model.
Physical Table Name	The CDA or DBA approved name(s) of the table(s) which corresponds to the entity. If the physical model has not been created, this is the table name as formed in accordance with the CMS standard naming conventions for the applicable DBMS. If the entity is marked as “Logical Only”, this value must be specified as “None”.

**Attribute Metadata Standards**

The following identifies and describes the required CMS Standard Attribute-Level Metadata that shall be captured for each Attribute in the Logical Data Model.

CMS Standard Attribute-Level Metadata	Attribute-Level Metadata Description
Attribute Derivation Text	Required for all derived attributes. The narrative explanation of any non-trivial logic used to transform information from one or more other attributes into the value for this attribute.
Attribute Requirement ID	<p>A reference to the requirement(s) or change request identifier(s) that justify the existence of the attribute in the model or changes to the attribute.</p> <p>Entry is required only when the attribute has been added as a result of a requirement that differs from the requirement entered in the “Entity Requirement ID” of the entity to which the attribute belongs. If the requirement for the attribute is the same as that of the entity, the “Attribute Requirement ID” is not required.</p> <p>Valid entry formats are as follows:</p> <ul style="list-style-type: none"> <li>(1) DOORS or requirement document Tracking ID, use BR-#### for business requirement, FR-##### for functional requirement, SR-#### for system requirements.</li> <li>(2) Change Requests, use CR-####.</li> <li>(3) Data change requests, use DR-#####.</li> <li>(4) Remedy Tickets, use RT_#####.</li> <li>(5) MAPD, use MAPD_####.</li> </ul>
Attribute Data Source Name	<p>The CMS database, external data feed, manual data entry process or software process from which the column takes its value. Valid entry formats are as follows:</p> <ul style="list-style-type: none"> <li>(1) CMS sources - database.table.column or filename.recordtype.field.</li> <li>(2) External system or document sources – These are (1) manual or automated system sources of data element names or (2) departmental document sources of data element names. The format is: org.specification.pubdate.record.subpart.field. Subpart is used where there are multiple record formats or levels of element structure, and may be omitted if there is only one flat format for the record specification. Example: This example shows a column data source field which is located three levels deep within an XML document "HealthLevelSeven.CDAL1R1.20061219.&lt;levelone_1.0.xsd&gt;.&lt;clinical_document_header&gt;.&lt;provider&gt;.&lt;type_cd&gt;".</li> <li>(3) Software processes - application name, service name, or “System Generated”.</li> <li>(4) Manual data entry processes - application.formname.fieldname , or “User-Supplied”</li> </ul> <p>If this level of detail is not available, indicate the source in terms of a user role, type of stakeholder, organization, business process, system, program, database, or file.</p>
Physical Column Name	<p>The CDA/DBA approved name(s) of the column(s) which correspond(s) to the attribute.</p> <p>If the physical model has not been created, this is the column name as formed in accordance with the CMS standard naming conventions for the applicable DBMS. If the attribute is marked as “Logical Only”, this value must be specified as “None”.</p>
Attribute Name	See: <a href="#">DM OP-012 Operating Procedure for Naming Data Attributes</a>

Data Management Operating Procedures and Guidelines

CMS Standard Attribute-Level Metadata	Attribute-Level Metadata Description
Attribute Definition	See: <a href="#">DM OP-010 Operating Procedure for Defining Data Attributes</a>
Attribute Alias Name	Any applicable alternate business name used to refer to the attribute.
Attribute CDA Standard Name	The CDA-approved name of the attribute, if different from the attribute's name in the model. This standard logical name must be used when creating a new data attribute to represent the identical concept in a new or existing model.
Attribute Domain Name	The name of the domain which defines the data type, default value and valid values of the attribute.
Attribute Logical Only Indicator	The indication of whether or not the Attribute has a corresponding column in a Physical data model type.
Attribute Valid Values	The list of values that the attribute can assume.
Attribute Required Indicator	An indication of whether or not the Attribute must assume a non-null value when an entity instance is created.
Attribute Primary Key Indicator	An indication of whether or not the attribute is part of the entity's primary identifier.
ANSI/ISO Standard Element Name	<p>The name of the corresponding data element defined in a national or international standard.</p> <p>The National/International Standard Element Name(s), preceded by the standard that corresponds to CMS attribute name. For example, the CMS Attribute Name= Beneficiary Birth Date would be reflected as follows:</p> <p><i>ANSI/ISO Standard Element Name HIPAA, Subscriber Date of Birth;C83,Person Date of Birth</i></p> <p>Note multiple standards are separated by semicolons.</p> <p>See: <a href="#">Data Model Tool Use Procedure</a></p>

## ***DM OP-032: Operating Procedure for Capturing the Standard Physical Data Model Metadata***

### *Rationale:*

Capturing a standard set of metadata associated with the Physical Data Model provides the necessary details to foster communication within and across projects.

### *Operating Procedure:*

The following metadata shall be captured for each Model, Table and Column in the PDM.

### **Model Metadata Standard**

The following identifies and describes the required CMS Standard Model-Level Metadata that shall be captured for each Physical Data Model.

<b>CMS Standard Model-Level Metadata</b>	<b>Model-Level Metadata Description</b>
Model Name	See: <a href="#">DM OP- 028 Operating Procedure for Naming and Defining Data Model</a>
Model Author	The first and last name of the Local or Central Data Architect as well as the company name of the person responsible for creating and/or maintaining the data model.
Model Definition	See: <a href="#">DM OP- 028 Operating Procedure for Naming and Defining Data Model</a>
Model Business Contact Name	The name of the organization and person who is responsible for approving the definitions in the model.
Model Central DA Name	The name of the person who is responsible for certifying the compliance of the model components according to CMS Data Administration Standards.
Model Create Date	The date recorded by Local DA for the original model.
Model Type	See: <a href="#">DM OP-029 Operating Procedure for Selecting Model Type.</a>
PDM Last Change Description	A brief narrative summarizing the nature of the changes resulting in the current version of the physical data model.
Model Central DBA Name	The name of the person who is responsible for certifying the compliance of database components according to CMS Database Administration Standards.
Model Local DBA Name	The name of the person who is responsible for the day to day operational support activities necessary to implement and maintain the database for a project.
PDM DA Signoff Date	The date when this version of the physical data model received its Central DA sign-off.

**Table Metadata Standard**

The following identifies and describes the required CMS Standard Table-Level Metadata that shall be captured for each table in the Physical Data Model.

CMS Standard Table-Level Metadata	Table-Level Metadata Description
Table Name	See: <a href="#">DM G-010 Guideline for Construction First Cut Physical Table or File Names from the Logical Data Model</a>
Table Comment	<p>The narrative explanation of the meaning of a row in the Table.</p> <p><i>Example: <b>TUS_CNTRCT - A binding agreement between CMS and an Organization that enables eligible Medicare beneficiaries to obtain medical services from the organization in exchange for monthly payments. Both current and historical information is retained.</b></i></p>
Physical Only Table Indicator	This property is <b>optional</b> and applies only if the table is a physical only table and is NOT documented in the Logical Model.
Logical Entity Equivalent Name	The CDA approved logical entity name that is equivalent to the table’s physical name. Required for tables marked as “Physical Only”.
Table Requirement ID	<p>A reference to the requirement(s) or change request identifier(s) that justify the existence of the Table in the model.</p> <p>Required for tables marked as “Physical Only”.</p> <p>Valid entry formats are as follows:</p> <p>(1) DOORS or requirement document Tracking ID use BR-#### for business requirement, FR-##### for functional requirement, SR-#### for system requirements.</p> <p>(2) Change requests use CR-####</p> <p>(3) Data change requests use DR-#####.</p> <p>(4) Remedy Tickets, use RT_#####.</p> <p>(5) For MAPD, use MAPD_####.</p>
Table Security Category Description	See: <a href="#">DM OP-021 Operating Procedure for Assigning Information Security Categories</a>
Logical Entity Name	The name of the corresponding entity or entities defined in the Logical Data Model which corresponds to this PDM.

**Column Metadata Standard**

The following identifies and describes the required CMS Standard Column-Level Metadata that shall be captured for each column in the Physical Data Model.

CMS Standard Column-Level Metadata	Column-Level Metadata Description
Column Name	See: <a href="#">DM G-011 Guideline for Constructing Physical Column or Element Names</a>
Column Comment	The narrative explanation of the meaning of an instance of the column. <i>Example: <b>CNTRCT_PBP_NAME – The name assigned by the contractor to the Plan Benefit Package.</b></i>
Column Domain Name	The name of the domain which defines the data type, default value and valid values of the column
Column Null Option	An indication of whether or not the Column must assume a non-null value when a row is created.
Physical Only Column Indicator	<b>Optional</b> and applies only if the column is a physical only column. Required for all columns in a physical only table.
Primary Key Indicator	An indication of whether or not the Column is part of the Table’s primary key.
Logical Attribute Equivalent Name	The CDA approved logical attribute name that is equivalent to the column’s physical name. Required for columns marked as “Physical Only”.
Logical Attribute Name	The name of the corresponding attribute, defined in the corresponding CDA approved Logical model , upon which this column's design was based.
Column Derivation Text	The narrative explanation of any non-trivial logic used to transform information from one or more other columns or fields into the value for this column.
Column Requirement ID	<p>A reference to the requirement or change request identifier(s) that justify the existence of the column in the model or changes to the column.</p> <p>Required for columns marked as “Physical Only”</p> <p>Entry is needed only when the column has been added as a result of a requirement that is different from the requirement entered in the “Table Requirement ID of the table that the column belongs to. If the requirement for the column is the same as that of the table, the “Column Requirement ID” is not required.</p> <p>Valid entry formats are as follows:</p> <ul style="list-style-type: none"> <li>(1) DOORS or requirement document Tracking ID use BR-#### for business requirement, FR-##### for functional requirement, SR-#### for system requirements.</li> <li>(2) Change requests use CR-####</li> <li>(3) Data change requests use DR-#####.</li> <li>(4) Remedy Tickets, use RT_#####.</li> <li>(5) MAPD, use MAPD_####.</li> </ul>

Data Management Operating Procedures and Guidelines

CMS Standard Column-Level Metadata	Column-Level Metadata Description
Column Data Source Name	<p>The CMS database, external data feed, manual data entry process or software process from which the column takes its value. Valid entry formats are as follows:</p> <p style="padding-left: 40px;">CMS - database.table.column or filename.recordtype.field.</p> <p style="padding-left: 40px;">External - org.specification.pubdate.record.subpart.field. Subpart is used where there are multiple record formats or levels of element structure, and may be omitted if there is only one flat format for the record specification.</p> <p>Example: This example shows a column data source field which is located three levels deep within an XML document  "HealthLevelSeven.CDAL1R1.20061219.&lt;levelone_1.0.xsd&gt;.&lt;clinical_document_header&gt;.&lt;provider&gt;.&lt;type_cd&gt;".</p> <p style="padding-left: 40px;">Software processes - application name, service name, or "System Generated".</p> <p style="padding-left: 40px;">Manual data entry processes - application.formname.fieldname , or "User-Supplied"</p> <p>If this level of detail is not available, indicate the source in terms of a user role, type of stakeholder, organization, business process, system, program, database, file or data exchange standard.</p>

### ***DM OP-033: Operating Procedure for Recording Data Model Changes***

#### *Rationale:*

A Project Logical Data Model (PLDM) will often undergo successive periods of change after the initial form of the model is reviewed and approved by Central Data Administration. These changes will be recorded on an ongoing basis by the Project Local DA. The benefits of recording such changes include:

1. Being able to present Central DA with a clear list of the changes that Central DA needs to review for a specific version of the Project Logical Data Model.
2. Providing the Project team with a concise historical record of when and why the valuable resource that the Project Logical Data Model constitutes was altered.

#### *Operating Procedure:*

1. Record changes to the PLDM objects using the template “Model Change Report Template” on [:http://www.cms.gov/DataAdmin/03\\_LogicalDataDesign.asp#TopOfPage](http://www.cms.gov/DataAdmin/03_LogicalDataDesign.asp#TopOfPage). The template contains instructions on how to fill out the fields.



## **Data Management Guidelines**

### ***DM G-001 Guideline for Assigning Data Analysts***

*Rationale:*

The quality of project data analysis and design deliverables relies on the capabilities of the Project Data Analyst.

*Guideline:*

Assign Project Data Analysts that have practical experience that is commensurate with the project's data analysis needs and the anticipated design complexity.

**Suggested Skill Set:**

1. A minimum of five years experience in Relational and/or Dimensional Modeling techniques.
2. Skills for compiling business requirements analysis and documentation.
3. Experience interfacing with health care professionals and business customers.
4. Skills and experiences with Medicare subject matter domains.
5. Knowledge of the health care insurance industry.
6. Experience with standard CMS data modeling tools.

## ***DM G-002 Guideline for Conducting the Logical Data Design Kickoff Meeting***

### ***Rationale:***

The purpose of this guideline is to encourage an understanding of participant roles and responsibilities, discuss major project goals, and confirm the level of agency sponsorship including project budget authorization. The meeting also provides an opportunity to discuss the data design process, the checkpoint reviews, and the location and scope of the data administration standards.

### ***Guideline:***

Follow the outlined agenda and exchange information as indicated.

#### **Agenda**

1. Introductions of team members and roles
2. Explain data design process
3. Identify level of service needed and available funding sources
4. Identify schedule requirements
5. Collect data requirements documents from Project Manager
6. Estimate time frame for follow-up response

***DM G-003 Guideline for Monitoring Schedule and Cost of Data Services***

*Rationale:*

Monitoring of logical data design costs is necessary to avoid unapproved costs and schedule overruns.

*Guideline:*

Monitor baseline estimates and report compliance on a periodic basis (e.g., bi-weekly) to the Project Manager.

***DM G-004 Guideline for Designating Representation Term and Data Type***

*Rationale:*

Attributes can be named and defined with additional consideration for how they are used and therefore, enhance their recognition and capability for storing business data values.

*Guideline:*

1. Suffix an attribute name with the appropriate *Representation Term* that fits with the anticipated category of data values to be stored.
2. There may be several physical data types that could be used to represent business data values for a domain. For example an integer data type (INTEGER, SMALLINT, TINYINT) could be used to store a domain whose valid values are integers between 1 and 20. Alternatively, a decimal data type could be used.
3. In selecting the data type, consider the manner in which the RDBMS of choice stores values and the application processing efficiency.
4. The Representation Term and data type must not conflict; for example an attribute ending with the representation term *Description* must not have a data type of *Decimal* assigned.
5. More examples of compatibility between Representation Term and data type are shown in the table below.

*CMS Standard Representation Terms and examples of associated Data Types*

TERM	ABBREVIATION	DEFINITION	Example data types
ADDRESS	ADR	Alphanumeric data which identifies street name and number, suite number, etc. May also be used for Post Office box, email address, or Internet address (URL).	String, Varchar, Char(50) Decimal (6,2)
AMOUNT	AMT	A number that is expressed as a currency denomination.	Currency
AUDIO	AUD	A type of Object that is capable of storing a stream of audible sounds.	String, Varchar
CODE	CD	A short character string used to replace a more descriptive value.	String, Char (50)
DATE	DT	A calendar date (month, day, and year) documenting when something happened.	Numeric
MESSAGE	MSG	A brief explanation or warning of a condition such as an edit validation error.	String, Varchar
QUANTITY	QTY	A type of numeric value that is a non-monetary measure of a dimensional extent, volume, etc.	Numeric, Integer
TIMESTAMP	TS	Date and time combined, documenting when something is recorded by a system.	Datetime

***DM G-005 Guideline for Constructing Physical Data Names***

DM G-005 has been deleted. It is replaced by *DM G-010 Guideline for Constructing Physical Table or File Names* and *DM G-011 Guideline for Constructing Physical Column or Element Names*.

***DM G-006 Standard for Assigning Date Formats***

***Rationale:***

For date and time attributes, Data Administration has adopted the International Standard for the representation of dates and times, ISO 8601. Its full reference number is ISO 8601: 2000 (E), and its title is "Data elements and interchange formats - Information interchange - Representation of dates and times." ISO 8601 describes a large number of date/time formats. The standard is based on the 24-hour timekeeping system. To reduce complexity, only the following date/time formats are being supported:

***Guideline:***

<b>Content Represented</b>	<b>Format</b>	<b>Example</b>
Complete date	YYYY-MM-DD	2003-10-08
Year and month	YYYY-MM	2003-10
Year	YYYY	2003
Complete time	hh:mm:ss	18:33:56
Hour and minute	hh:mm	18:33
Hour	hh	18

DA has approved these formats for logical design; this standard is NOT intended to constrain physical implementation.

## ***DM G-007 Guideline for Conducting the Data Architecture Review***

### ***Rationale:***

The purpose of this guideline is to obtain an agreement on the sources of project entity level data as early in the data design process as possible. Ideally, an early agreement will give all the impacted parties the time to accommodate the new data needs.

### ***Guideline:***

Follow the outlined agenda and exchange information as indicated.

### ***Required Materials:***

1. A draft Project Logical Data Model, which, at a minimum, shows the project entities. (Detailed level attributes are not required for the Architecture Review.)
2. A Data Source Plan
3. An Impact Analysis Report

### ***Agenda:***

1. Introduce the project principals and briefly discuss the project's purpose.
2. Based on the Data Source Plan:
  - a. Review the EDM candidate entities that are being proposed for reuse;
  - b. Review the non-EDM data sources that are being proposed for reuse; and
  - c. Discuss / justify the need for any new entities.
3. Based on the Impact Analysis Report:
  - a. Discuss the pros and cons for production application systems that might be impacted by the project application's shared use of existing data structures.
4. Request timely approval of the Data Source Plan (or an alternative recommendation) from the responsible Data Architect(s) and Data Steward(s).

### ***DM G-008 Guideline for Mapping Data Requirement Sources***

*Rationale:*

Mapping is performed to validate attribute-level completeness of the Project Logical Data Model to meet business needs, to assure that attributes have a source of their values defined and to confirm any requirements for aggregated or derived data..

Each attribute in a Project Logical Data Model must contribute to at least one business requirement. Requirement statements in the standard requirements management tool are cross-referenced to each attribute in the data modeling tool.

The origin of an attribute's content can be:

- User input
- Database column(s) and containing table or view
- Data exchange standard, message type and field(s)
- Software process

This information may be documented from an initial load perspective, and from an on-going operations perspective.

*Guideline:*

Record the corresponding requirement information for each attribute using the Attribute UDP named Attribute Requirement ID

Record the data source information for each attribute using the Attribute UDP named Attribute Data Source Name.

If applicable, document the transformation algorithm used to compute the attribute's value in the Attribute UDP named Attribute Derivation Text.

The Project Data Analysts' role in developing, or supporting documentation of this information may vary.

## ***DM G-009 Guideline for Preparing the Project Logical Data Model for Physical Design***

### *Rationale:*

The logical data modeler and physical data modeler may or may not be the same individual. In either case, the logical data modeler should anticipate what the physical data modeler needs to know in order to develop an efficient physical data model and capture these suggestions as comments in the entity definition

### *Guideline:*

1. Consider the performance of record retrieval based on anticipated record volume and the caching, indexing, and clustering facilities available in the anticipated DBMS.
2. Consider the mandated enterprise data security settings and any conflicts with the proposed business data functional requirements. For example, whether data will be distributed to multiple platforms or outside of firewalls.
3. Make notes of additional concerns and insights, which can later be considered in physical data design.
4. Do not modify the project logical data model to reflect physical design compromises such as denormalization to roll down supertype attributes into subtype entities. Information pertinent to physical design considerations are to be recorded as comments.
5. Provide the first-cut physical names using the approved abbreviated terms and construction rules defined in *DM G-010 Guideline for Constructing Physical Table and File Names* or *DM G-011 Guideline for Constructing Physical Column or Element Names*

***DM G-010 Guideline for Constructing Physical Table or Files Names***

*Rationale:*

Physical data names represent the technical implementation of important business concepts. Therefore, it is important that the names of physical data can be easily recognized and associated with the business concept. Physical data names representing system operational requirements are important to database administration. Therefore, it is also important that physical-only objects can be easily tracked, understood, and compared.

*Guiding Principles:*

1. Physical data name lengths must comply with the name length limitation established by Central DBA standards for the target database management system (DBMS).
2. Approved acronyms must be used in place of standard abbreviations. For example, Centers for Medicare and Medicaid Services must be abbreviated as CMS, not CTR\_MDCR\_MDCD\_SRV.

*General Steps to Attain First Cut Physical Table or File Names*

Step 1. (For logical/physical models the first two steps are done automatically. For physical only models they must be done manually)

- Translate logical data model to physical data model by using CMS Standard Terms and Abbreviation List.
- Replace space between terms with an underscore.
- Analyze the length of table names and proceed to step 2 if they exceed the DBMS table name length limitation. Otherwise proceed to step 3.

Step 2.

- Eliminate the least important terms until the target length is attained.

Step 3.

- Ensure table names are unique within the physical data model.

*Examples:*

Logical Entity Name	Physical Table Name
ORGANIZATION ENROLLMENT	ORG_ENRLMT
ORGANIZATION	ORG
BENEFICIARY DUAL STATUS	BENE_DUAL_STUS

## *DM G-011 Guideline for Constructing Physical Column or Element Names*

### *Rationale:*

Physical data names represent the technical implementation of important business concepts. Therefore, it is important that the names of physical data can be easily recognized and associated with the business concept. Physical data names representing system operational requirements are important to database administration. Therefore, it is also important that physical-only objects can be easily tracked, understood, and compared.

### *Guiding Principles:*

1. Physical data name lengths must comply with the name length limitation established by Central DBA standards for the target database management system (DBMS).
2. Approved acronyms must be used in place of standard abbreviations. For example, Centers for Medicare and Medicaid Services must be abbreviated as CMS, not CTR\_MDCR\_MDCD\_SRV.
3. Every column and element name must have a representation term.

### *General Steps to Attain First Cut Physical Column or Element Names*

Step 1. (For logical/physical models the first two steps are done automatically. For physical only models they must be done manually)

- Translate logical data model to physical data model by using CMS Standard Terms and Abbreviation List.
- Replace space between terms with an underscore.
- If the last term in the physical column or element name is an acronym, append the approved abbreviation of the representation term to the physical name. For example, translate logical attribute Beneficiary Identification Code to physical column BIC. Append abbreviation for Code (representation term) to the column or element name resulting in BIC\_CD.
- Analyze the length of column names and proceed to step 2 if they exceed the DBMS table name length limitation. Otherwise proceed to step 3.

Step 2.

- Eliminate the least important terms until the target length is attained.
- Representation term may not be removed.

Step 3.

- Ensure column names are unique within the corresponding table.

Step 4.

- For an LDM/PDM, update the physical name in the PDM if the default name has been changed
- For separate LDM and PDM, there are two situations:
  - If no PDM has yet been created then the use of the ERwin “Derive a new model” feature in DM G-010 above to create an ERwin logical\physical model from the separate LDM will result in the physical view of this ERwin logical\physical model having names that can serve as the first cut column names.
  - If the project team has already produced a separate PDM, then there is no need for the Central DA to provide first cut physical column names.
- For separate LDM and PDM, update Physical Column Name UDP for attribute in LDM

*Examples:*

Logical Attribute Name	Physical Column Name
ORGANIZATION ENROLLMENT DATE	ORG_ENRLMT_DT
PERSON LAST NAME	PRSN_LAST_NAME
PROVIDER NATIONAL PROVIDER IDENTIFIER	PRVDR_NPI_ID

## *DM G-012 Guidelines for Assigning Data Planning Participants*

### *Rationale:*

Business management and IT management must work together to create an Enterprise Architecture (EA) for CMS business functions.

### *Guideline:*

Assign CMS personnel and responsibilities as described below.

1. CMS Enterprise Architect (FEA liaison)  
Maintains the enterprise architecture to ensure that business processes and information systems are aligned.
2. Enterprise Data Architect  
Identifies data design and integration issues.
3. Enterprise Process Architect  
Identifies process design and integration issues.
4. Business Owners / Partners  
Provide business information requirements and security category specifications.
5. Data Stewards  
Ensure that all processes enforce data quality.

## *DM G-013 Guidelines for Conducting Enterprise Data work sessions*

### *Rationale:*

This is a guideline for conducting Information Technology Systems Planning work sessions related to Enterprise Data Architecture.

### *Guideline:*

Follow the outlined agenda and exchange information as indicated.

1. Schedule design sessions with Business Intelligence Council and meeting facility.
2. Conduct orientation and training for design session participants.
3. Prepare the materials, room, and software aids.
4. Customize the design session agenda.
5. Conduct the kickoff meeting.
6. Agenda
  - a) Introductions of BIC team members and roles
  - b) Designate the BIC Chairperson.
  - c) Establish the need for alignment between business functions with data architecture.
  - d) Select team members for subsequent detailed definition component and break-out sessions.
  - e) Explain how the EA project fits in the strategic direction of the agency.
  - f) Identify level of service needed and available funding sources
  - g) Explain the definition design process:
  - h) Confirm the general requirements for the business functions areas affected, including the scope, goals and objectives.
  - i) List the main processes within each affected business function area.
  - j) Develop blueprints for and descriptions of the process and enterprise entities in use (the as-is).
  - k) Identify the information needed for supporting the processes and depict the enterprise entities in a logical data model (the to-be).
  - l) Document expected individual contributions.
  - m) Identify schedule requirements.
  - n) Estimate the timeframes for follow-up responses.
  - o) Publish work session minutes.

## ***DM G-014 Guideline for Creating a Business Process Diagram***

### *Rationale:*

This is a guideline for development of a Business Process Diagram.

### *Guideline:*

1. Define the process scope.
  - a. Describe what is being done.
  - b. Tie the process to an organization and business application.
2. Set the level of detail
  - a. Scope the diagram to represent a major business process at *Level 0*. (This is equivalent to a FEA-BRM chartered function such as *Maintain Medicare Beneficiary Enrollment*).
  - b. Decompose the *Level 0* process to 8-10 *Level 1* sub-processes.
  - c. If needed to show all business recognized tasks, decompose respective *Level 1* sub-processes to *Level 2* sub-processes.
3. Present the processes in graphical format.
  - a. It is not necessary to show sequence.
  - b. Is not necessary to express conditions that initiate a process.
  - c. Show organization interaction (who does what).
  - d. Show information exchanges (what is needed to do this).
  - e. Identify the FEA- BRM alignment.

## *DM G-015 Guidelines for Defining a Subject Area*

### *Rationale:*

Data architecture is more helpful to business personnel when it is organized in a manner that speaks to business purposes. That is the objective of a Subject Area: to serve as the broadest classification for a business line's data.

### *Guideline:*

Create a Subject Area by the following criteria:

1. Name and define a Subject Area according to business line. Use a work session to allow business users to discuss their processes and come up with an overall name for a grouping of enterprise entities. Good subject area names usually reflect a major business grouping of data. Occasionally, a subject area name might be the name of a particular business activity that an organization keeps information about, for instance *Beneficiary Enrollment*.
2. Enterprise Entities are exclusive to a single Subject Area. Do not duplicate enterprise entity or subject area names.
3. All *Enterprise Entities* must have an assigned *Subject Area*.
4. Data Stewards should be stakeholders from all parts of the Subject Area business line. Try to assign Data Stewards from the organization unit that is as close as possible to the initial point of data capture of the most important Enterprise Entities in the *Subject Area*.
5. If a business change means that certain business functions end, determine whether Enterprise Entities can be removed from Subject Area(s) and the Enterprise Data Model.

***DM G-016A Guidelines for Analyzing and Approving Standard Terms***

***Rationale:***

A consistent use of business terms leads to a common vocabulary and understanding of business data. Therefore, it is important that the business taxonomy be managed and used appropriately to label data entities and attributes.

***Guideline:***

If terms in the *Standard Term List* are inadequate for naming a data entity or attribute, a request may be made to add a new standard term to the list.

New terms are analyzed according to the following criteria:

1. All standard terms are to be atomic, containing no special characters e.g., hyphens, slashes, or underscores.
2. Terms are accepted in singular form e.g. “Vendor” is acceptable “Vendors” is not.
3. Slang or buzz words are not acceptable e.g. “Zippy”, or “Buy-in”.
4. Terms are accepted in the present tense e.g., “Interviewed” is accepted as “Interview”.
5. Terms are accepted in their root form e.g., the acceptable form of “Rotation” is “Rotate”.
6. When a word can be expressed in different forms to indicate possession or direction, multiple forms of a term can be accepted e.g., “Payee”, “Payer”, and “Payment”.
7. Connectors are not accepted e.g., “Of”, “By”, “Per”, “Or”, “And”.
8. Avoid addition of synonyms to the list, except in those instances where the synonym’s use is justified for business communication. For example, when the term “Responsibility” is an existing standard term, resist adding the synonym “Duty”.
9. Avoid negative terms such as “Non-participant”, “Unqualified”, “Ineligible”, etc. Rationale: the use of negative terms invariably leads to confusing names; also, they have the potential for creating double negative conditions.

Term approval criteria	Example of data name, using negative term - “Ineligible” versus positive term - “Eligible”	Example of data value	Likely business interpretations or questions
Avoid negative terms	Person Ineligibility Switch	“N”	The “N” value means that the person is eligible, or does it?
	Person Ineligibility Switch	“Y”	Does this mean the person is eligible?
Allow positive terms	Person Eligibility Switch	“N”	Clearly, the person is not eligible.
	Person Eligibility Switch	“Y”	Clearly, the person is eligible.

10. Explicit criteria for Object Class Terms (compliance with ISO/IEC 11179 – 5): An Object Class Term is a noun that represents an important entity classes or entity subclasses in an enterprise data subject area. Examples of Object Class Terms: Account, Agreement, Budget, Employee, Location, Procurement, Supplier
11. Explicit criteria for Property Terms (compliance with ISO/IEC 11179 – 5): A Property Term expresses a discrete category. Examples of Property Terms: City, Gender, Spouse, Report, Revenue

Explicit criteria for Qualifier Terms (compliance with ISO/IEC 11179 – 5): A Qualifier Term can be used to further modify or specify an Object Class Term or Property Term. Examples of Qualifier Terms: Minimum, Net, Last, Temporary

12. Explicit criteria for Representation Terms (compliance with ISO/IEC 11179 – 5).
13. Representation Term describes the valid form of data content. Add new terms only when a new form of data can not be accommodated or expressed by the data types associated with existing Representation Terms. (See DM G-004 Standard for Designating Representation Term and Data Type)
14. Examples of Representation Terms: Code, Identifier, Number, Date.

## ***DM G-016B Guidelines for Abbreviating a Standard Term***

### ***Rationale:***

Consistent abbreviations of business terms support the naming of physical data artifacts. Therefore, it is important that an abbreviation of a term can be easily recognized and readily associated with the original business term.

### ***Guideline:***

If there is no common or industry-wide abbreviation for a word, use the following guidelines to create an abbreviation. (Example abbreviated terms are shown in brackets.)

1. Words that are four letters or less are not abbreviated unless there is a widely-recognized and common abbreviation available.

*Example:* AGE [AGE]

*There are exceptions such as:* DATE [DT], CODE [CD]

2. Condense all double constants to a single constant.

*Example:* Enrollment [ENRL]

3. Eliminate vowels except: a.) when it is a leading vowel; b.) when the vowel is needed for clarity; c.) when the vowel is needed to provide uniqueness; d.) when a vowel is part of a commonly used abbreviation.

*Example of a:* Employee [EMPY]

*Example of b:* Exposition [EXPO]

*Example of c:* Library [LIB] and Liability [LIAB] (An "A" is added to distinguish Liability's abbreviation).

*Example d:* Product [PROD]

4. Abbreviations are limited to eight characters. Abbreviations of 4 characters are preferred.

*Discouraged example:* RESPONSIBILITY [RSPNSBLTY]

*Suggested example:* RESPONSIBILITY [RESP]

5. Abbreviations must be unique in order to support clarity of usage.

*First term example:* RESPONSIBILITY [RESP]

*Second term example:* RESPONSE [RSPN]

6. An abbreviation may not form a word. If the abbreviation forms a word, replace letters to make the abbreviation not be a word while maintaining uniqueness.

*Prohibited example:* RETENTION [RENT]

*Suggested example:* RETENTION [RTNT]

7. Abbreviated words that are widely known and consistently used in the agency should be defined as an acronym.

*Example:* Health Service Area [HSA]

8. Numbers are not abbreviated, and can be attached to a word.

*First example:* 1993 [1993]

*Second example:* ALTERNATIVE 1 [ALTN1]

## ***DM G-017 Guidelines for Reference Tables***

### *Rationale:*

The purpose of a *reference table* is: 1) to categorize, classify, qualify, and constrain other data; 2) to provide a method of coding and decoding; and 3) to ensure that only valid values are entered into a database.

### *Guideline:*

1. Reference tables must be free of defects.
2. Reference tables are part of physical data design.
3. Reference tables may be included on the logical data model at the discretion of the project data modeler and the data architect.

The use of reference tables is favored under the following circumstances:

4. If the number of allowed values is at least twenty (20);
5. If the set of allowed values is subject to change or grow; and
6. When it is simpler to change values in a table rather than updating program logic or adding a physical check constraint.

Reference tables tend to follow a standard format:

7. *Primary identifier, Code, Full Name, and a Description.*

***DM G-018 Guidelines for Managing Data Values through Physical Constraints***

***Rationale:***

The integrity of data values in data resources is important in providing an organization with quality information. Database Management Systems provide various methods that can automatically check and ensure that values comply within an established range.

***Guideline:***

The following list describes types of constraints.

1. *Primary Key Constraint* - The *primary key* of a table holds a value that uniquely identifies each row in the table. It is defined when a table is created or altered. A table may contain only one *primary key*.
2. *Foreign Key Constraint* - A *foreign key* is a column or combination of columns that form a relationship between the data in two tables. The relationship is created between two tables by adding the column or columns that hold one table's primary key values to like column or columns in the other table. The second table holds the *foreign key*.
3. *Unique Constraint* - A *unique* constraint ensures that no duplicate values are entered in specific columns that do not participate in a *primary key*.
4. *Check Constraint* - A *check* constraint enforces domain value integrity by limiting the values that are permitted in a column. Like *foreign key* constraints, *check* constraints control the values that are placed in a column. In contrast, *check* constraints limit the valid values to a *declared* list of values.
5. *Default Constraint* -The *default* constraint determines the value to be stored when no value is explicitly assigned.
6. *Null Constraint* -This determines whether a column accepts a NULL value.

*Integrity constraints* attempt to enforce data consistency and prevent accidental damage to the database during updates. Use the table below as a guideline for when to apply alternative constraint types.

Managerial Circumstance	Constraint Type			
	Assertions	Check Constraints	Foreign Key (FK) Constraints	Triggers
A stable set of permitted values		√		
Occasional change to the set of permitted values		√	√	
The data value is dependent on variable conditions.	√			√

**DM G-019 Guideline for Modeling Supertypes and Subtypes**

*Rationale:*

It may be helpful to partition an entity type into subtype entities when: it helps business communication to show a subtype entity with a more specialized meaning; or when a subtype entity requires different integrity treatment.

*Guideline:*

1. A supertype or subtype is named according to entity naming standards.
2. A subtype has at least one attribute or one relationship that is not in the supertype.
3. Sometimes it is simpler to combine a supertype and subtype within an entity. However when subtypes are represented within a single entity, the attributes specific to the subtype are defined as *optional* (nulls allowed). This permits the entity to exist when the subtype does not, as illustrated in the left-side example in the following table.

<i>Example: combined Supertype /Subtype</i>	<i>Example: partitioned Supertype / Subtype</i>
<p><i>Employee</i></p> <ul style="list-style-type: none"> <li><i>Employee Number(required)</i></li> <li><i>Employee Name(required)</i></li> <li><i>Employee Job Code(required)</i></li> <li><i>Employee Street Address(required)</i></li> <li><i>Employee Temporary Type Code(required)</i></li> <li><i>Employee Temporary Company Name (optional)</i></li> </ul>	<p><i>Employee</i></p> <ul style="list-style-type: none"> <li><i>Employee Number(required)</i></li> <li><i>Employee Name(required)</i></li> <li><i>Employee Job Code(required)</i></li> <li><i>Employee Permanent Address (required)</i></li> <li><i>Employee Temporary</i></li> <li><i>Employee Number(required)</i></li> <li><i>Employee Temporary Company Name (required)</i></li> </ul>

4. Note the attribute *Employee Temporary Type Code*; this attribute distinguishes whether the record holds data for a temporary employee or not and is referred to as a *type code* or a *discriminator*. *Discriminators* are needed when subtypes are generalized into the supertype.
5. Balance the need for subtyping by looking at business rules: create a subtype when there is a distinct need for subtype *referential integrity*; otherwise, generalize the subtype within the supertype.
6. Observe normalization rules in order that the Logical Data Model transforms appropriately into a Physical Data Model.

## ***DM G-020 Guideline for Using ERwin Complete Compare to Create a Brand New Data Model***

### *Rationale:*

To manage data models as an enterprise information resource, it is necessary for the model metadata to be collected in a consistent form. The Tool Use standard defines how standard UDPs will be used as the mechanism to promote such consistency. This guideline describes the procedure for incorporating the standard data model UDPs into a brand new model.

### *Guideline:*

When a modeler is tasked with creating a brand-new data model, the templates can be used as a start point to facilitate the introduction of the standard UDPs. The modeler will simply:

1. Open up the relevant logical\physical, logical or physical data model template
2. Do a “save as” in order to save the template under the name desired for the brand new data model.
3. The modeler can now open up the model just created from the template.
4. In the main subject is a block which should be filled in by the modeler with the appropriate identifying information.
5. This model will initially contain absolutely no entities, tables, etc. But as a consequence of having used the template as a start point, every new object that is subsequently added (manually or by using "Complete Compare") to the new model will immediately have the correct standard UDPs, ready for whatever entry is relevant. The modeler is thus spared the effort of having to add the UDPs to a new data model.

***DM G-021 Guideline for Naming Reference Entities***

*Rationale:*

A reference entity contains a valid value code attribute as the primary key and a valid value description attribute as a non-key attribute. A good data entity name is meaningful and is based on the respective entity definition. It provides the first source of information about the entity’s purpose and contents.

*Guideline:*

1. Use REFERENCE as the last term in the entity name.
2. Select an object class term that best describes the type of data. (i.e., do not use DATA as an object class term).

<b><i>Recommended Name</i></b>	<b><i>Not Recommended Name</i></b>
Beneficiary Sex Reference	Beneficiary Sex
Claim Admission Type Reference	Data Admission Type Table
Claim Frequency Reference	Claim Frequency Code
Product BETOS Reference	Data BETOS Table