

Client Logo

<Client Name>
Data Model Standard

Version 0.1

<Date>



Table of Contents

Revision History	3
Objective	4
Scope	4
Definitions	4
Standard Statement	8
Enterprise Logical Data Model.....	8
Data Modeling Notation.....	8
Data Model Tools	9
Data Model Object Naming Conventions.....	9
Logical Model	9
Logical Data Model Requirements.....	10
Physical Model.....	11
Physical Data Model Requirements	11
Validation of Data Models.....	12
Roles and Responsibilities	12
Review Cycle	12
Effective Date.....	12

Revision History

Version	Date	Description	Author

Objective

The objective of this standard is to define Information Management Structure and Design naming standards

Scope

This standard covers Logical Data Models and Physical Data Models

Definitions

Term	Definition
Alternate Key	An attribute or set of attributes that uniquely identify an entity but are not the primary key. For example, an Identification Number is an alternate key for the entity VEHICLE OPERATOR LICENSE. The primary key is License Number.
Attribute	<p>A property of an entity, able to contain a characteristic (value).</p> <p>As defined in a logical data model, an attribute becomes a column and can be used as a primary key in the physical data model. Such an attribute consists of three words:</p> <ul style="list-style-type: none">• Prime word• Qualifier word• Class word <p>For example, Account Number is an attribute of the entity ACCOUNT.</p>
Logical Data Model	A Logical Data Model that reflects consistent definitions and business rules at the enterprise level. This data model is the starting point for all project data models.
Cardinality	A property of a relationship between two entities that specifies whether there is one instance, more than one instance or a specific number of instances of one entity related to one instance of the other entity. The three generally used cardinality values are one-to-one, one-to-many and many-to-many.
Column	The physical implementation of an attribute in a database management system.

Term	Definition
Conceptual Data Model	A data model that shows only the fundamental entities and relationships. Conceptual models usually show only a minimum set of attributes, if any.
Database	A structured collection of data, usually managed by database management system software.
Data Mart	A type of data warehouse designed primarily to address a specific function or department's needs.
Data Warehouse	A system for storing, retrieving and managing large amounts of any type of data.
Data Type	A property of data attributes or columns. Logical data types classify the general form of the data (e.g., character, number, date, etc.), and physical data types specify particular formats for a technology (e.g., varchar, float, integer, etc.).
Dimensional Data Model	A type of Physical Data Model that may be used in data mart and data warehouse design.
Dimension (Table)	A table that provides descriptive or lookup information in a Dimensional Data Model.
Domain	A logical or physical format for data that allows for standardization and reuse. For example, a domain for dollar amounts provides for up to 15 digits to the left of the decimal point and two digits to the right.
Entity	A person, place or thing, essentially a noun that describes a concept of interest to the enterprise. For example, ACCOUNT, ORGANIZATION, TRANSACTION. Entities are logical constructs.
Exclusive Subtype	A subtype that does not allow a supertype instance to belong to another subtype simultaneously. For example, an instance of the entity EXEMPT EMPLOYEE cannot also be an instance of the entity NONEXEMPT EMPLOYEE. Both are exclusive subtypes of the entity EMPLOYEE.
Fact Table	A table that has columns that contain the quantities and amounts that provide the detailed data in a Dimensional Data Model.
Foreign Key	An attribute or set of attributes in an entity that reflect a relationship to another entity. These foreign key attribute(s) are the primary key in the other entity (ies).

Term	Definition
Identifying Relationship	A relationship between two entities in which the primary key of one of the entities forms part of the primary key of the other entity in the relationship. An identifying relationship shows that one entity is fully dependent on the other.
Inclusive Subtype	A subtype that allows an instance of a supertype to belong to another subtype simultaneously. For example, the subtype entity WATER-BASED AIRCRAFT and the subtype entity LAND-BASED AIRCRAFT are inclusive subtypes of the supertype entity AIRCRAFT because an amphibious aircraft can belong to both subtypes simultaneously.
Logical Data Model	A data model that shows data and data relationships without regard to a specific implementing technology. A Logical Data Model includes fully defined subject areas, entities and attributes and reflects business requirements only. A Logical Data Model cannot be implemented directly. A Logical Data Model should always be normalized to at least third normal form.
Logical Data Model Objects	Those items that are the components of a Logical Data Model. These include data entities, data attributes, data relationships, domains, data types and valid values.
Model Seeding	The process by which a new data model is created from selected entities and relationships in a pre-existing data model. Model seeding provides consistency, traceability and reuse between models.
Non-identifying Relationship	A relationship between two entities that does not result in the primary key of one entity forming part of the primary key of the other.
Orphan Entity	An entity in a Logical Data Model that has no relationship to any other entity.
Physical Data Model	A data model that reflects decisions that meet the requirements of a specific implementation technology, including those related to usage and performance. A Physical Data Model is implemented directly.

Term	Definition
Physical Data Models Objects	A data model that reflects decisions that meet the requirements of a specific implementation technology, including those related to usage and performance. A Physical Data Model is directly implemented. Within the Physical Data Model class, there are the Relational Physical Data Models and Dimensional Physical Data Models.
Primary Key	The attribute or set of attributes that uniquely identifies an entity. Tables may also have a column or set of columns that act as a <i>physical</i> primary key.
Relational Data Model	A Physical Data Model that is intended to be used in the design of an operational, transactional application. Contrast with a Dimensional Data Model.
Relationship	An association between two or more entities. The relationship reflects a business rule about the association. Relationships usually indicate how often the relationship may hold and whether it is mandatory or not.
Role Name	A descriptive word added to an attribute name, placed before the original name, when that attribute is used more than once within an entity because of relationships. For instance, CAMPAIGN STRATEGY has a role-named attribute of Parent Strategy Identifier to distinguish the role of this Strategy Identifier.
Subject Area	A subset of entities that has been selected from the main data model primarily used to provide focus on a particular business data requirement or function.
Subtype Discriminator	An attribute in a supertype that contains data values that serve to identify the possible subtypes. For example, Employee Type Code, with the values E for exempt and N for non-exempt, is the subtype discriminator for EMPLOYEE.
Subtype	An entity that represents a subset of another entity (called the supertype). For example, EXEMPT EMPLOYEE and NONEXEMPT EMPLOYEE are both subtypes of EMPLOYEE. Subtypes inherit all of the attributes and relationships of the supertype but have either unique attributes or relationships or both.
Supertype	An entity that has one or more subtypes.

Term	Definition
Table	The physical implementation of an entity in a database management system.
Transform	A mechanism, typically found in the data-modeling tool, that provides documentation and traceability for changes made to a Logical Data Model in converting it to a Physical Data Model.
Valid Value	An item of data subject to typing rules, codeset rules, and validation rules such as, but not limited to, conforming to predefined constraints or semantic restrictions; belonging to a set, controlled vocabulary, or enumeration list; subject to an exclusion rule; based on a derivation rule; or falling within a range or set of ranges.

Standard Statement

Data modeling standards are established so that Conceptual, Logical and Physical Data Models are designed and structured according to the same rules and guidelines and are applied on a consistent basis.

Failure to comply with the Data Modeling Standard shall result in sanctions and/or actions determined by the Chief Information Officer (CIO)

Enterprise Logical Data Model

The starting point for all data models shall be the Enterprise Data Model (EDM).

All conceptual, logical and physical data models will be stored in an approved enterprise software, as documented in the current version of the Enterprise Technology Standards List, with the controls established to ensure appropriate version control, recovery, access and maintenance.

Data Modeling Notation

Data models shall be designed and illustrated following an approved notation or methodology. Applying an approved notation or methodology on a consistent basis leads to easily read and understood data models of high integrity.

Data Model Tools

Only approved tools shall be used when designing, building, maintaining and storing data models.

Approved tools for designing, building, maintaining and storing data models are listed in the current version of the Enterprise Technology Standards List.

Data Model Object Naming Conventions

Each data model object, when used in a data model, shall be named in accordance with the Data Model Naming Guideline.

Logical Model

Model Names:

A standard for model naming should be established.

Subject Areas

Subject area names should be a singular noun phrase. Every logical model should incorporate and use subject areas. Each subject areas should contain a definition.

Entities

- An entity is something about which the business needs to record data;
- Entity names should be a singular noun phrase that is not abbreviated. Examples of entity names are: customer, product, time, location, vendor, etc.
- Each entity should contain a definition and example.

Attributes

- An attribute is a named indivisible element of information about an entity;
- Attributes describe the characteristics or properties of an entity and the information we want to know about an entity'
- Attributes should be singular noun phrases that are not abbreviated but are in business language terms; and
- Each attribute should contain a definition and example.

Relationships

- Relationships must be defined for logical models;
- Relationships should define business rules by showing the interrelationships between two entities, not one indirectly defined through other entities;
- Relationships are shown as lines between the entities. The relationship name is composed of the entity names, the parent-to-child verb phrase based on the relationship type, and the cardinality e.g.

- Customer places one or many orders.
- One and only one customer places order

Volumetric Storage Parameters

- Volumetric data should be captured and stored in the Logical model showing:
 - Initial Volume
 - Growth per month; and
 - Maximum Volume

Logical Data Model Requirements

When a Logical Data Model is completed, it shall have the following characteristics:

- No orphan entities;
- All entities are defined;
- All attributes are defined;
- Every entity shall have its primary key identified;
- Alternate keys shall be identified as “Alternate Keys” to aid in data access analysis and physical design;
- All entities, except some associative entities, shall have at least one “Non-primary-key” attribute;
- All subtypes shall be identified as “inclusive” or “exclusive.”
- Shall have a subtype discriminator;
- Every attribute shall be identified, as either “required” or “not required” and the corresponding relationship shall be specified as to whether nulls are either “allowed or not allowed.”
- Every relationship shall be specified as either “identifying” or “non-identifying.”
- Every relationship specifies whether nulls are allowed;
- A role name shall be specified if there is more than one valid relationship involving the same entities or if the relationship is recursive;
- Every attribute shall have a domain assigned;
- All known valid values shall be documented; and
- Attributes shall have a data type.

Physical Model

Tables

Table names should be derived from logical entity names. Each should table contain a definition and example.

Columns Names and Attribute Names

Column names should be derived from logical attribute names. Each attribute should contain a definition and example.

Views

- Views not directly accessed by users should follow standards for table and column names;
- Views directly accessed by users should be named using natural language, i.e., fully spelled out names;
- When the name exceeds a length supported by the RDBMS, apply standard abbreviations to reduce the name length; and
- The recommended technique for shortening names is to apply standard abbreviations/acronyms from right to left.

Physical Data Model Requirements

When a Physical Data Model is completed, it shall have the following characteristics:

- Data models shall have tables;
- Data models shall have columns;
- Data models shall have primary keys;
- Data models shall have foreign keys;
- If an entire table is physical only, any relationships between that table and another table shall also be marked as physical only; and
- The domain is defined for each attribute.

Validation of Data Models

A validation process shall be documented and approved for every data model to ensure the model is complete, consistent and compliant with the standard statements.

Validation for logical data models shall include the following:

- Entities without attributes;
- Attributes without definitions;
- Entities without primary keys;
- Entities without definitions; and
- Entities without relationships.

Validation for physical data models shall include the following:

- Indexes without columns;
- Columns without comments;
- Tables without primary keys;
- Redundant indexes;
- Tables without columns;
- Columns with default datatypes; and
- Columns with different datatypes

Roles and Responsibilities

Data Modeler is accountable to the technical lead for ensuring that logical and physical data models conform to naming standards

Review Cycle

This Data Modeling Standard must be reviewed annually

Effective Date

The Data Modeling Standard is effective <date>. All new data models developed must adhere to the elements as defined in this Standard.