



DELAWARE STATE-WIDE INFORMATION TECHNOLOGY AND ARCHITECTURE STANDARDS

Standard ID:	IN-DAT-003
Title:	Data Modeling
Domain:	Information
Discipline:	Modeling
Effective Date:	5/15/2013
Revision no.:	1
Original date:	11/10/2011

I. Authority, Applicability and Purpose

- A. Authority** – Title 29, Chapter 90C provides broad statutory authority to the Department of Technology and Information to implement statewide and interagency technology solutions, policy, standards and guidelines for the State of Delaware's technology infrastructure. "Technology" means computing and telecommunications systems, their supporting infrastructure and interconnectivity used to acquire, transport, process, analyze, store and disseminate information or data electronically. The term "technology" includes systems and equipment associated with e-government and Internet initiatives.
- B. Applicability** – Applies to all State of Delaware communications and computing resources. DTI is an Executive Branch Agency and has no authority over the customers in Legislative and Judicial Branches, as well as School Districts, and other Federal and Local Government entities that use these resources. However, all users, including these entities, must agree to abide by all policies, standards promulgated by DTI as a condition of funding, access and continued use of these resources.
- C. Purpose** – Due to the importance of the information managed by the State's technology solutions, it is necessary to establish common guidelines for Data Modeling. This document provides approaches and best practices for Data Modeling.

II. Scope

- A. State of Delaware** – All communications and computing resources involved with data owned by the State of Delaware
- B. Areas Covered** – This standard covers all data and data modeling technologies whether they were developed in-house or purchased as a complete solution.
- C. Environments** – This standard addresses all environments that contain State of Delaware data, managed by State of Delaware Data Stewards.

These standards are adopted by the Department of Technology and Information (DTI), through the Technology and Architecture Standards Committee (TASC), and are applicable to all Information Technology use throughout the State of Delaware. Any questions or comments should be directed to dti_tasc@state.de.us.



DELAWARE STATE-WIDE INFORMATION TECHNOLOGY AND ARCHITECTURE STANDARDS

III. Process

- A. **Adoption** – These standards have been adopted by the Department of Technology and Information (DTI) through the Technology and Architecture Standards Committee (TASC) and are applicable to all Information Technology use throughout the state of Delaware.
- B. **Revision** – Technology is constantly evolving; therefore the standards will need to be regularly reviewed. It is the intent of the TASC to review this standard annually. The TASC is open to suggestions and comments from knowledgeable individuals within the state, although we ask that they be channeled through your Information Resource Manager (IRM).
- C. **Contractors** – Contractors or other third parties are required to comply with these standards when proposing technology solutions to DTI or other state entities. Failure to do so could result in rejection by the Delaware Technology Investment Council. For further guidance, or to seek review of a component that is not rated below, contact the TASC at dti_tasc@state.de.us.
- D. **Implementation responsibility** – DTI and/or the organization's technical staff will implement this standard during the course of normal business activities, including business case review, architectural review, project execution and the design, development, or support of systems.
- E. **Enforcement** – DTI will enforce this standard during the course of normal business activities, including business case and architectural review of proposed projects and during the design, development, or support of systems. This standard may also be enforced by others during the course of their normal business activities, including audits and design reviews.
- F. **Contact us** – Any questions or comments should be directed to dti_tasc@state.de.us.

IV. Definitions/Declarations

A. Definitions

- 1. **Attribute** – An attribute is another name for a column in a database schema.
 - 2. **Conformed Dimension:** A conformed dimension is a set of data attributes that have been physically implemented in different databases using the same structure, attributes, domain values, definitions and concepts.
 - 3. **Data Dictionary** – It contains the non-technical (business terminology) definitions of fields.
 - 4. **Data Modeling** – Method used to define and analyze data and the requirements needed to support the business process. The final product will be a true and current representation of the production database. The Data Model is a living document and will change in response to the business. The Data Model also, defines the structure and relationship between the data elements. The three main types of Data Models are Conceptual Data Model, Logical Data Model and Physical Data Model. The preferred sequence for doing Data Modeling is:
 - ✓ Conceptual Data Model
 - ✓ Logical Data Model
 - ✓ Physical Data Model
- **Conceptual Data Model** – This Data Model describes data requirements from a business point of view without the burden of technical details. Models at this level are about understanding the data requirements of the business.

These standards are adopted by the Department of Technology and Information (DTI), through the Technology and Architecture Standards Committee (TASC), and are applicable to all Information Technology use throughout the State of Delaware. Any questions or comments should be directed to dti_tasc@state.de.us.



DELAWARE STATE-WIDE INFORMATION TECHNOLOGY AND ARCHITECTURE STANDARDS

- **Logical Data Model** – This Data Model refines the conceptual models by documenting the entities, their attributes and their relationships. These models are technology oriented designs, although they are database-independent.
 - **Physical Data Model** – This Data Model represents the detailed specification of what is physically implemented using specific technology. Physical design considerations include performance, size and growth, availability, recovery from failure, and use of specific technology features.
5. **Data provenance:** Data Provenance refers to the process of tracing and recording the origins of data and its movement.
 6. **Dimensional Modeling:** Dimensional Modeling is a design technique for databases intended to support end-user queries. It is oriented around business understandability and performance. It is a de-normalized design that adds fact tables (central tables containing measurements) related to dimensional tables (contains core data, which is being measured).
 7. **Entity:** Entity (table) represents a concept, object or thing which is relevant to the application being developed and about which information needs to be held by the application.
 8. **Meta-model** - The meta-model is the cornerstone in the metadata repository architecture as it defines the physical data model that stores the metadata. There are industry standards for meta-models that enable greater degrees of tool metadata sharing and interoperability. These standards don't guarantee interchange without any effort, but do give you a starter set for entities and attributes to include the metamodel.
 9. **Metadata** – Data that describes the data. A Metadata record consists of a set of elements that describe the characteristics of an information asset or resource. The more detailed the metadata (especially the business explanations) the better it will be for system implementation, usage and maintenance.

Consistency in the metadata is necessary to keep information organized. Consistent terminology helps communicate metadata, and it helps applications process the metadata. The Categories of Metadata are:

- **Analytical** - Analytic Metadata describes the derivations and display of reporting environments. Primary sources of analytic metadata include OLAP and reporting packages metadata environments.
- **Business** - This category of metadata defines in a business context the information that the data provides. Examples of business metadata are business attribute names, business attribute definitions, business attribute valid values, data quality rules, data models and business rules. Primary sources of business metadata include logical data models and data quality.
- **Navigational** – Navigational metadata describes the data linkage and data movement within the environments. Examples of navigational metadata are derived fields, business hierarchies, source columns & fields, transformations, data quality checkpoints, target columns & fields and source & target locations.



DELAWARE STATE-WIDE INFORMATION TECHNOLOGY AND ARCHITECTURE STANDARDS

- **Operational** – Operational metadata describes the data integration applications and jobs through statistics giving a full technical view of the environment. Examples of Operational metadata include jobs statistics and data quality check results. Primary sources of Operational metadata include data integration job logs and data quality checks.
 - **Structural** - Structural metadata provides the description of data within the IT infrastructure. For example, where the data is located, the names under which it can be accessed, what kinds of data types are being stored, data lineage and data integration within Client's IT environment. Examples for Structural Metadata are:
 - Databases / File groups
 - Tables / Views / Files
 - Keys
 - Indices
 - Columns/fields
 - Source columns/fields
 - Target columns/fields
10. **Normalized Model** -- In the design of a relational database management system (RDBMS), the process of organizing data to minimize redundancy is called normalization. The objective is to isolate data so that additions, deletions, and modifications of a field can be made in just one table and then propagated through the rest of the database via the defined relationships.
 11. **Relational Model:** Organizes the data elements into tables and creates relationships between the tables. The model is the blueprint to the database and is used for managing database objects, writing reports and queries, and writing application code.
 12. **Star Schema:** The star schema consists of one or more fact tables referencing dimension tables. The fact table(s) contains the measurements (summaries, aggregations, etc) and the dimension tables contain the business objects (people, places, etc).
 13. **Subtype:** Subtypes are entities that have the same attributes as a more generalized entity. Internal organizations, external organizations, internal contracts, and external contracts are examples of subtypes.
 14. **Supertype:** Supertype is the more general entity that contains the same attributes as subtype. Organizations and contracts are examples of supertypes.



B. Declarations

Data Model Tool must:

- Be able to export the metadata to common format (XML, Text and Microsoft Excel) which would help in repository sharing.
- Have a central repository.
- Be able to search objects (tables, columns, constraints) across the entire model.
- Generate code for multiple types of databases.
- Able to import or export data models created or consumed by other data model tools.
- Provide the ability to re-use objects across models and automatically create linkage for object use.
- Be able to provide impact analysis within and across models.
- Provide the ability to create connections, mappings, and dependencies between models.
- Be able to export data models to a common viewable format where users can see the data model without the Data Model tool.
- Have an entry in the central repository that consists of its model type, name, definition, and applicable characteristics such as a data type for database columns.

Data Model must:

- Conform to the State's Data Naming standard(s) when it is published.
- Have Data Terms that are submitted to the State's data dictionary.
- Be created for Conceptual Data Model and Physical Data Model.
- Have Conceptual Data Model approved by business before other models are started, created or modified.
- Focus the conceptual data model on the business needs and contain business vocabulary.
- Contain the following when a logical data model is created:
 - Entities and Attributes
 - Primary key attributes and non-key attributes must be defined
 - Relationships must be defined.
 - Include the data steward, data classification and data sharing rules.
- Be broken out by subject areas



DELAWARE STATE-WIDE INFORMATION TECHNOLOGY AND ARCHITECTURE STANDARDS

- Have Non-Project specific reference tables in a separate subject area. This will help in re-use.
- Physical Data Model must be database specific and must define database implementation tasks such as database stored procedures, triggers, database size, backup and partitioning etc.
- Contain the following for Physical Data Models:
 - Schema, Views, Tables, Columns, Indexes, Constraints
 - Tables must be defined. Any table partitioning required for the implementation should be defined.
 - Column must have a definition and must specify data type, length, optionality and keys.
 - Table and Column name must be sized to fit the requirements of the target DBMS tool.
 - Table and Column name cannot have blanks or spaces and cannot start with a number.
 - Document the range or list of values for each column. This can be enforced using constraints or relationships of other tables.
 - Define alternate keys that will enhance performance by supporting common paths.
 - Performance improvements can be achieved through clustered index, caching and index optimization based on the target DBMS platform.
 - Define security requirements for attributes with data classification higher than public and plan for implementation of security policies.
 - Referential integrity rules must be defined for constraints on a foreign key for updates or deletion.
- At minimum, a conceptual data models or data dictionary must be submitted to DTI once the application design has been finalized or prior to production implementation of the application. DTI will preserve the data model in a central repository and apply the data model to the enterprise data model based on the fit. The data model or data dictionary submitted to DTI must include at least the following items:
 - Entity names and descriptions
 - Entity relationships and descriptions
 - Attribute names, descriptions, data type, and length
 - Primary identifier for each entity
- Data dictionaries or data models for vendor solutions must contain the core entities or tables that house state-owned business data.



DELAWARE STATE-WIDE INFORMATION TECHNOLOGY AND ARCHITECTURE STANDARDS

Data Model(s) may:

- Generate the Logical Data Model from the Conceptual Data Model and related business and functional requirements.
- Generate the Physical Data Model from the Logical Data Model and related business, functional and technical requirements
- Be created for legacy data stores especially logical data models
- Have notation for physical data models and preferred to be expressed in UML notation.
- Have notation for logical data models and are preferred to be expressed in E/R notation.
- Physical data model can be implemented as relational, dimensional or normalized.
- Physical data model can have a certain amount of denormalization which will help in performance gain by the downstream system.
- Physical data model can contain computed values if they provide efficiency for the retrieval of data by avoiding joins.
- Table and Column names should be singular for physical data models.
- Each Column name in a Physical data model should contain all of the elements of the logical attribute from which it was derived, but should be abbreviated to fit within the maximum length.
- Be arranged for ease in visual understanding

Metadata tool must:

- Have a central repository.
- Be able to pull all related metadata stored by tools and other applications to the central repository.
- Store the following information:
 - metamodel and data store requirements
 - metamodel requirements and metadata sources
 - Load mechanism, CRUD (Create Read Update Delete) and administration requirements
 - Roles and Responsibilities for use
 - Capture history and versioning
- Have an Audit trail. (Security, Metadata access .. etc)
- Have an online interface for viewing definitions
- Have an interface for exporting Metadata into a common format (XML, Text, Excel)



DELAWARE STATE-WIDE INFORMATION TECHNOLOGY AND ARCHITECTURE STANDARDS

MetaData must:

- Be adopted by State Organizations / School Districts / Quasi State Organizations through Data Stewards.
- Contain the business rules and Lineage of data movement.
- Contain Data provenance
- Contain Data Model which has specification of data structures and business rules and also they act as the source of business and technical metadata.
- Store each data element associated with a degree of privacy and whether its use is subject to limitations.
- Be revisited on an annual basis by the respective organizational data stewards to ensure the details of terms stored reflect the business or purpose it was intended for.

MetaData may:

- Have technical metadata created through the design tools.
 - Have operational metadata created by any executable.
-



DELAWARE STATE-WIDE INFORMATION TECHNOLOGY AND ARCHITECTURE STANDARDS

V. Definition of Ratings

Individual components within a Standard will be rated in one of the following categories. COMPONENT RATING	USAGE NOTES
<ul style="list-style-type: none"> • STANDARD – DTI offers internal support and has arranged for external vendor support as well (where applicable). DTI believes the component is robust and can be expected to enjoy a useful life of 5+ years from the Effective Date. 	<p>These components can be used without explicit DTI approval for both new projects and enhancement of existing systems.</p> <p><i>(1) Note the useful life concern for the “Acceptable” rating.</i></p>
<ul style="list-style-type: none"> • ACCEPTABLE – DTI offers internal support and has arranged for external vendor support as well (where applicable). DTI believes the component is stable, but has a useful life ⁽¹⁾ of less than 5 years from the Effective Date. 	
<ul style="list-style-type: none"> • EMERGING – DTI considers the component to be a likely candidate for future classification as STANDARD or ACCEPTABLE within the state pending further investigation. 	<p>These components must be explicitly approved by DTI for all projects. They must not be used for minor enhancement and system maintenance without explicit DTI approval.</p>
<ul style="list-style-type: none"> • DECLINING – Deprecated – DTI considers the component to be a likely candidate to have support discontinued in the near future. A deprecated element is one becoming invalid or obsolete. 	
<ul style="list-style-type: none"> • LIMITED SUPPORT – DTI has limited or no internal support capability for the component; or has no arrangement for vendor support for the product. Users must arrange for adequate overall support of the component through their own efforts. 	
<ul style="list-style-type: none"> • NOT SUPPORTED BY DTI – DTI offers no internal support and has no arrangement for vendor support. Users must arrange for all support of the component through their own efforts. 	
<ul style="list-style-type: none"> • DISCONTINUE – For reasons of overall risk, product support, high TCO, or other issues, the use of this technology is discouraged. All current instances of this technology should have a plan developed for its retirement. DTI expects to work aggressively with the users of such technologies to devise a collaborative plan. 	<p>No waiver requests for new solutions with this component rating will be considered.</p>
<ul style="list-style-type: none"> • DISALLOWED – DTI declares the component to be unacceptable for use and will actively intervene to disallow its use when discovered. 	<p>No waiver requests for new solutions with this component rating will be considered.</p>

These standards are adopted by the Department of Technology and Information (DTI), through the Technology and Architecture Standards Committee (TASC), and are applicable to all Information Technology use throughout the State of Delaware. Any questions or comments should be directed to dti_tasc@state.de.us.



DELAWARE STATE-WIDE INFORMATION TECHNOLOGY AND ARCHITECTURE STANDARDS

- A. Applicability of Ratings** – The ratings and usage notes are intended to encourage technology decisions to move toward components that enjoy the full support of DTI. However, acknowledging that mass replacement of lower rated components is not feasible, DTI will allow continued maintenance, enhancement, and possibly limited new development using these components. In making such determinations, DTI may require that the requestor demonstrate that they have adequate support arrangements in place.
- B. Missing Components** – No conclusions should be inferred if a specific component is not listed. Instead, contact the TASC to obtain further information.

VI. Component Assessments

All implementations of the following components much adhere to the [State of Delaware's standards and policies](#). Of particular note are [State of Delaware Information Security Policy](#), [State of Delaware Data Classification Policy](#) and [Retention Schedules](#).

#	Component	Rating	Comments
<u>1</u>			
a)	Sybase PowerDesigner	Standard	
b)	CA ERwin	Not Supported By DTI	
c)	IBM Rational Data Architect	Discontinue	
d)	Visio	Discontinue	