By Steve Hoberman

Steve Hoberman & Associates, LLC

The Teradata Communications Industry Logical Data Model

Overview and Application

Sponsored by:

TERADATA
Raising Intelligence

Table of Contents

Introduction

After graduating college, I was hired as a data modeler for a telecommunications research company. Eight other modelers and I worked heads down full time with the single goal of building the enterprise data model for the telecommunications industry. It was an extremely challenging project, with the ambitious plan of using this model to drive an entire new suite of operational applications to run the business. After ten years (I was on the team for two of these years), our model had more than 3,000 entities which covered an entire corridor outside our offices.

Not every communications company today has ten years and nine full-time resources budgeted to create the *big picture* for their organization. We still need this big picture though, to reap the rewards of getting the right information to the right people at the right time. A single representation of Customer, for example, enables graceful growth of operational information and the building blocks for powerful business intelligence (BI) applications.

A well-understood big picture of the organization needs to be captured and communicated in the form of a model. A model is a set of symbols and text used to make a complex landscape easier to grasp. The world around us is full of obstacles that can overwhelm our senses and make it very challenging to focus only on the relevant information needed to make intelligent decisions. A complex geographic *landscape* is made understandable via a model called a map. A complex *information* landscape is made understandable via a data model. A data model uses symbols and text to help developers and analysts understand a set of data elements and the corresponding business rules better. In addition, every model has a defined scope. A map might be limited to New York City or represent the big picture in the form of

"An industry data model is a prebuilt data model that captures how an organization in a particular industry works or should work."

a globe. Likewise, a data model can represent a specific functional area, such as order processing, or it can represent the big picture in the form of an enterprise data model (EDM).

An EDM is a subject-oriented and integrated data model describing all of the data produced and consumed across an entire organization. Subject-oriented means the concepts on a data model fit together as the CEO sees the company, as opposed to how individual functional or department heads see their view of the company. There is one person who can play many roles including possibly being both a Customer and an Employee. Integration goes hand in hand with subject orientation and implies a single version of the truth along with a mapping back to the chaotic real world. For example, if a person's last name resides in ten applications within an organization, the integrated EDM would show Individual Last Name only once, and in addition, capture the mapping back to these ten applications, such as the person's last name as a Customer and an Employee.

There are resource and skill challenges with creating and maintaining an EDM, and therefore, organizations are increasingly purchasing starter EDMs in the form of industry data models instead of reinventing the wheel. An industry data model is a prebuilt data model that captures how an organization in a particular industry works or should work. Teradata Corporation offers eight industry data models called industry Logical Data Models (iLDMs):

- Teradata® Manufacturing Logical Data Model (MLDM)
- Teradata Communications Logical Data Model (CLDM)
- Teradata Financial Services Logical Data Model (FS-LDM)
- Teradata Healthcare Logical Data Model (HCLDM)
- Teradata Media and Entertainment Logical Data Model (M&E LDM)
- Teradata Retail Logical Data Model (RLDM)
- Teradata Travel and Hospitality Logical Data Model (T&H LDM)
- Teradata Transportation Logical Data Model (TLDM)

In the Teradata white paper titled, *Leveraging the Industry Logical Data Model*, I provided an overview to the Enterprise Data Model and the Teradata iLDMs. In this white paper, I will go into detail about the Teradata Communications industry LDM (CLDM). Specifically, this paper provides an overview of the Teradata CLDM and a scenario illustrating how the CLDM can be leveraged. The goal of this paper is to increase your awareness of how the CLDM helps organizations obtain a big picture more quickly and more accurately than building an EDM from scratch, thus permitting your organization to answer complex strategic and tactical business questions faster and more accurately.

Teradata CLDM Overview

The Teradata CLDM captures how a general communications organization works. It provides the big picture for a communications organization, containing nine broad subject areas: Network, Network Activity, Event, Financial Management, Party, Master (i.e., location-centric concepts, such as telephone number, site, and address), Promotion, Offer, and Finance. I've studied industry models that were extremely generic, and therefore only contained a handful of generic entities (e.g., Party). These generic models appear elegant yet require extremely complex mappings to the real source system to produce any value. The CLDM does contain a handful of these generic concepts (e.g., Event), yet these generic concepts are used to link more granular and concrete parts of the business together (e.g., a service order, billing statement adjustment, and power outage are all Events), and even to link different iLDMs together (e.g., Event appears in several of the iLDMs). Due to the details provided in the CLDM, the source system mapping becomes more manageable. The current version of the CLDM is extremely robust, containing more than 1,500 entities, but these numbers – and model features – are continuously updated through new releases.

The Teradata CLDM is a living, breathing view of the communications business. This model provides a holistic view of customer management, revenue assurance, network asset management, product management, and financial management functions. Teradata Professional Services consultants work directly with clients in the field and provide feedback for model changes and enhancements to the Teradata Product Manager who then captures these new requirements for potential addition in the next CLDM release. In September 2007, for example, Release 10.0 included adding new functionality, such as Call Center Management and Retail Sales Transactions. Each iteration of the Teradata CLDM results in CLDM customers benefiting from the enhancement suggestions from many earlier CLDM implementations.

The Teradata CLDM exists in an ERwin® Data Modeler file. ERwin Data Modeler is one of the more popular data modeling tools that supports reports for viewing and printing the models and their meta data. In addition, the CLDM documentation includes both hard copy and pdf files of three books. These include a reference guide common across all Teradata iLDMs, a reference guide specific to the CLDM, and a set of nine appendices. Appendix A walks through a number of scenarios (e.g., New Customer), and shows which entities from the CLDM are needed for each scenario step. Appendices B, C, and D contain nearly 100 sample business questions, along with a handy reference showing which parts of the CLDM are needed to answer each question. Appendix E contains customization guidelines, such as how to make generic structures such as Party more specific. Appendix F contains the entity definitions, and Appendix G contains the attribute definitions. Appendix H contains common abbreviations and acronyms (Know what USSGL stands for?). Appendix I is a collection of semantic models that illustrates how different parts of the CLDM can be combined to provide common reporting needs such as understanding customer churn.

The Teradata CLDM has several very important characteristics:

Logical

A logical data model is a business solution for a specific set of business requirements. If a requirement is to capture Call Detail Record (CDR) information, the logical data model would contain

the data elements and business rules around the call. It is completely independent of both application and technology, built using the process of normalization. Normalization ensures all data elements are correctly assigned to entities based on their dependency on a primary key ("Every data element depends upon the key, the whole key, and nothing but the key.").

Extensible

The CLDM contains the common information that companies share within an industry, and, therefore, it is meant to be a jumpstart toward creating a complete solution for a company. Most companies use the CLDM as a starter model, and add new structures, remove existing structures, and enrich the provided definitions to make it more meaningful to the organization.

Abstract

The CLDM contains a fair amount of abstraction. Abstraction means combining like things together under generic terms, such as Event and Party, to facilitate integration and to gracefully handle future requirements. For example, a subscriber buys a new handset at a retail outlet and calls the call center two weeks later. Treating the handset purchase and the call center call as events allows one to understand the dependencies between these events, answering queries such as "Is SMS usage greater after a call center call?" Also, the CLDM can easily accommodate a new type of Event, as well as *connect* with other iLDMs that also use the Event concept. This allows for greater commonality within and across the iLDMs. All industries have Events, whether they are service outages in the communications industry, campaign solicitations in the banking industry, or bookings in the travel industry.

Global

The structures and terms on the CLDM are designed for international use, and are not just U.S.-based. For example, the term 'postal code' is chosen over 'zip code' and 'territory' instead of 'state'. This facilitates communication on global projects and mappings back to global source systems, such as ERP systems including SAP® R/3®.

Standard

The CLDM follows best practice naming standards, including the use of class words based on the International Standards Organization (ISO) 11179 meta data standard. A class word is the last part of a data element name that represents the high-level category in which the data element belongs. Examples of class words are name, code, identifier, date, quantity, and amount. For example, the class word for Person Last Name is 'Name'.

Digestible

The CLDM is sectioned into subject areas. Subjects are neatly captured in separate views, and color is used to distinguish each subject area making it easier to digest the larger models. In addition, there are certain subject areas that are common across the iLDMs, such as Party and Financial Management. These subject areas have a common core in each iLDM, and then are extended where appropriate within each of the models.

Teradata CLDM Scenario

RU-There (RUT), a medium-sized communications company in the Midwest, has been consistently losing market share over the past five years. The CFO of RUT is at a loss to explain the specific reasons behind the declining market share other than to relate it to increased competition. Without understanding the cause, it's difficult to come up with a turnaround plan. For example, should the focus be on reducing costs, introducing new products, or reducing customer churn?

RUT has grown rapidly by purchasing smaller providers, and data integration has always taken a back seat. Many siloed operational and reporting systems make it nearly impossible to answer any business questions that cross departments or business functions, including those of importance to the CFO. It is for this reason the CIO has been in disguise. The fake mustache is starting to cause some face irritation though. Therefore, the CIO has initiated a project to produce RUT's enterprise data model to use as a foundation to build integrated applications that can answer important questions, such as those asked by the CFO. Can the enterprise data model be implemented before the CIO's disguise is discovered?

The Approach

Jamie Jitterbug, a highly skilled data analyst in RUT's enterprise data management team, is responsible for building RUT's EDM. She built four data models: white board conceptual data model (CDM), enterprise CDM, enterprise LDM, and an enterprise physical data model (PDM). The white board CDM was built without any reference to the CLDM. The enterprise LDM was built using the CLDM in four different roles, which are discussed later in this paper. The enterprise PDM was built based completely on the enterprise LDM. Figure 1 summarizes each of these models, and the following sections will provide the details along with examples.

White Board Conceptual Data Model

Jamie organized a series of meetings with business analysts, functional analysts, and department managers with a goal of creating a single, high-level view of the organization. She met with groups of one to five individuals and built their views of the organization using whiteboards and flipcharts. For those individuals who preferred not to see data models, Jamie worked with them to jointly create a listing of key concepts and their definitions. The finished model had severe integration issues as you might expect. Sets of entities were not related to each other, and there were many cases where the same concept had two or more definitions, and similar concepts had completely different names and rules. This is actually a very good thing because it documents the integration issues, and acknowledging the problem is a prerequisite to solving the problem. Jamie called this initial model the white board CDM because most of it was created in partnership with the business managers standing at white boards and flipcharts. It represented each business area in their terms.

Model	Purpose	Built all at once or incrementally
White board CDM	Captures the current understanding of the business on one piece of paper.	All at once
Enterprise CDM	Captures a proposed integrated view of the business on one piece of paper.	All at once
Enterprise LDM	Captures a cross-functional, objective, and detailed view of business data.	Incremental
Enterprise PDM	Captures a detailed view of the business data taking into account the constraints of the database management system and user queries.	Incremental

Figure 1. Types of EDMs.

The concept of Offering will be used to illustrate the four different types of models in this section. Offering is an entity within the Order subject area. There were three different definitions of Offering identified in the white board CDM, as shown in Figure 2.

An **offering** is any package we sell directly to our customers. It can contain one or more product components. [from the sales department manager]

An **offering** is anything we can market to make money. [from the marketing department manager]

An **offering** is anything sold by our organization. [from an accounting department representative]

Figure 2. Three different definitions of Offering.

This white board CDM had more than 200 entities. It was built all at once using a top down approach. A top down approach is one where the model is built purely from the business perspective and not from an existing systems perspective.

Enterprise Conceptual Data Model

The Teradata CLDM comes with a Subject Area Model that contains nine key concepts and their relationships for the communications industry. It captures the concepts and relationships across the CLDM's broad subject areas. For example, more than 100 entities, including Offering, Product, and Incentive, are represented by just the single Offer entity on the Teradata Subject Area Model. Figure 3 shows the Teradata Subject Area Model.

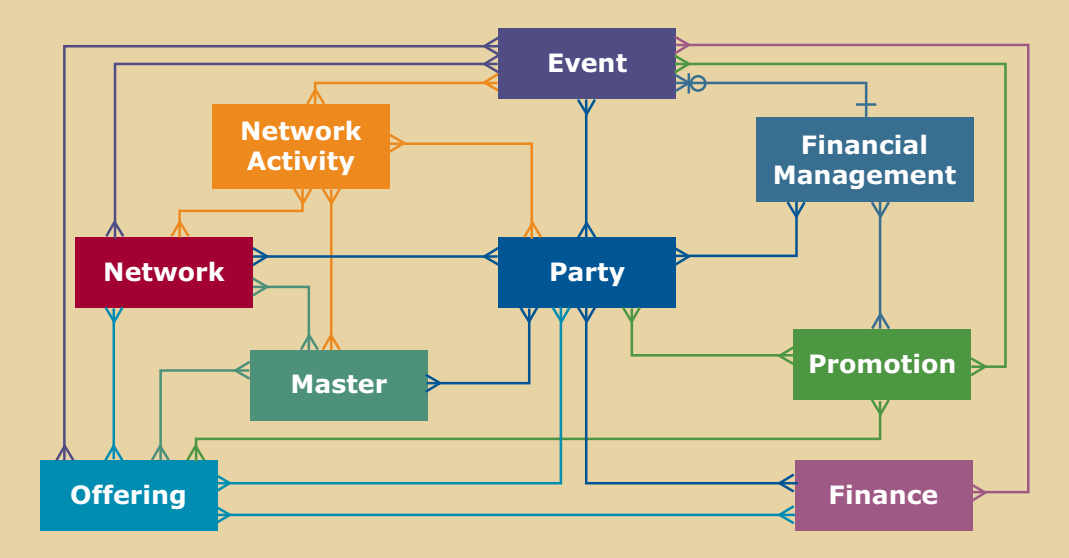

Figure 3. Teradata Subject Area Model.

The Teradata Subject Area Model allows an organization to achieve a high-level, big picture of the organization without getting overwhelmed by jumping straight into a complex logical design. The colors used on this model for each subject area are also used for all of the entities within each subject area. For example, all of the Event entities in the CLDM are purple to be consistent with Event in the Subject Area Model. Jamie took a first pass at fitting the white board CDM into the Teradata Subject Area Model.

After speculating how the pieces might fit together, she organized a second series of meetings. These meetings took place in groups of 10 to 15 individuals, and Jamie purposely invited people with very different views about the same concepts. She showed them the CDM retrofitted with each of their views and encouraged open communication so that when the meeting was over, there would be either agreement on the model or issues that would have to be reconciled.

Figure 4 contains the portion of the Enterprise CDM after the terminology and definitions surrounding the term offering were resolved.

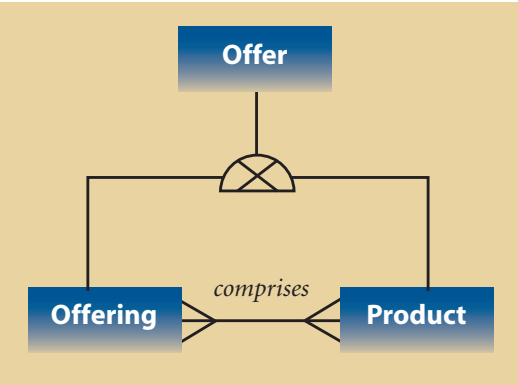

Figure 4. Subset of Enterprise CDM.

The semicircle with an X in the middle is a subtype symbol. It identifies a grouping entity (in this case OFFER) called a supertype, as well as those entities sharing common data elements and relationships (in this case OFFERING and PRODUCT) called the subtypes. OFFER already existed in the Teradata Subject Area Model, and OFFERING and PRODUCT are logical entities in the CLDM that were brought up into the Enterprise CDM to help clarify the differences between these concepts.

The definition for OFFER in the Teradata Subject Area Model is:

The Offer subject area provides information about the Products or Product Packages (Offerings) that are sold by the Communications Service Provider. Offerings may contain one or more Products. Offerings have characterizing Descriptors, such as Cost per Minute of Usage or Monthly Recurring Charge.

The definition for OFFERING in the Teradata CLDM is:

A grouping of one or more Products and the corresponding price plan, rate structure, or unit charge that may be marketed by the Communications Service Provider for the purpose of generating revenue. An offering is what a customer purchases or subscribes to.

Examples:

- Rate plan local cellular service, voice mail, wireless phone, and caller ID for $30.00 per month
- Basic residential line with call waiting, caller ID, and call forwarding for $19.95 per month
- Nokia 6100 Wireless Tri-Mode telephone for $200.00

The definition for PRODUCT in the Teradata CLDM is:

A telecommunications service, commodity, piece of equipment, or special service that may be marketed for revenue-generating purposes. A PRODUCT could be owned or developed by someone other than the CSP. The Party's role with a PRODUCT is represented in PARTY PRODUCT.

Examples of partner/competitor products include:

- Cable wholesale partner: DSL
- IP wholesale partner: IP Bandwidth
- Long distance carrier: Interconnect transfer
- Roaming partner: Roaming access
- Content provider: Content items
- Competitor products: Rate plans for rating analysis

The sales department manager's definition for offering (Figure 2) most closely matched the definition for OFFERING; the marketing department manager's definition most closely matched the definition of PRODUCT; and the accounting department representative's definition most closely matched the concept of OFFER. In each case, the definitions were expanded to include examples specific to RUT.

Jamie's CDM was regarded as a large success within both business and IT circles. Jamie credits the success to first attempting to understand the organization and then leveraging the Teradata model. Even business people with very strong viewpoints found it easier to adopt the CLDM terminology rather than get into a win/lose debate with colleagues from different departments.

Enterprise Logical Data Model

As you might expect, the Enterprise Logical Data Model (LDM) required more effort than the prior two models. It had more detail and required the most discussions to resolve the integration issues. Note that some of the integration issues remained unresolved yet well documented. Version 1 of RUT's Enterprise LDM contained more than 900 entities and 2,300 data elements. It was built using a hybrid approach. A hybrid approach means it was built from both a top down and bottom up perspective. Top down is driven from the business requirements, which takes the form of the Enterprise CDM, and bottom up means start with the existing systems environment.

Teradata CLDM Roles

An industry data model can play up to four different roles within an organization: blueprint, template, encyclopedia, and invisible. These are described below in order of decreasing reliance on the CLDM (e.g., blueprint requires the most reliance on the CLDM and invisible the least). The degree of reliance is determined by available modeling resources and knowledge of a particular business process.

Blueprint (The industry data model is *the* model)

The CLDM contains the concept of an Analytical Model within the Party subject area. The definition of Analytical Model is: "Describes a process used to predict, cluster, or classify information. Typically used in data mining and knowledge discovery. Examples: Customer Scoring and Segmentation, a model that describes the propensity of a customer to engage in a particular activity, etc." Analytical Model is a concept that the organization has not even considered relevant, yet after understanding its potential value for predicting future market share and profitability, they decided to add it to their EDM. This involved adding more than a dozen new entities to their EDM exactly as they appeared in the CLDM, including the actual Analytical Model entity.

Template (The industry data model is an integration point)

The CLDM concept of Offering becomes an important integration point for the company. Each of the Offering data elements from the source systems was mapped into CLDM data elements. A sample mapping appears in Figure 5.

Note that this mapping is overly simplified, as usually there can be complex transformation rules, as well as other types of meta data that need to be reconciled, such as format, granularity, and nullability.

Many integration battles are quickly defused using the CLDM, because instead of win/lose definition debates among business managers, it becomes a mapping exercise where both parties agree on a single, external, unbiased view.

Source			CLDM	
Source system	**Table or file**	**Data element**	**Entity**	**Data element**
XYZ	MERCHANDISE	Effective_Date	OFFERING	Offering_Start_Dt
ABC	PACKAGE	Begin_Date		Offering_Start_Dt
X3000	PLAN	Start_Date		Offering_Start_Dt

Figure 5. Data element sample mapping.

Encyclopedia (The industry data model is referenced where needed)

There is a need within the organization to better understand product capabilities and relate these capabilities to the actual Product. The CLDM provides a comprehensive data model containing product capability concepts. Jamie researched this area in the CLDM and was able to understand the data and rules behind the model so she can add these concepts to the existing EDM. In some cases, terms, rules, and definitions from the CLDM needed to be changed to fit the existing EDM. David Schoeff, Teradata Principal Consultant, compares this approach with how someone would use an encyclopedia: "There can be a substantial amount of modeling needed to build an organization's EDM, and the iLDMs can serve as a valuable reference to save some modeling time and reduce risk by ensuring all concepts are present on the model."

Invisible (The industry data model is not consulted)

The CLDM is not used at all. The Address area is extremely well modeled within RUT and has been rigorously maintained for the past five years. For this area, the CLDM was not consulted at all. Parts of the CLDM that were used and contained address information were connected to RUT's existing address structures.

Enterprise Physical Data Model

The Enterprise Physical Data Model (PDM) was built incrementally on a project-by-project basis. An in-depth business questions analysis was performed, and sets of business questions were bundled into project deliverables. Jamie found it challenging to extract questions from the business folks. Luckily, the CLDM came with approximately one hundred typical business questions, and she used this list as a brainstorming technique with the business managers to agree on a set of common questions. In fact, one business question from the CLDM list became the scope for an entire data mart: "How many churned Customers have called the Customer Care Center 30 days prior to canceling service?"

Smooth Sailing Ahead

All of RUT's future operational and business intelligence applications relied on the EDM as a starting point for design. When each application data model was considered complete, a review took place to identify possible EDM changes as a result of this application model. So each application starts with the EDM, and then contributes new ideas back to the EDM. This keeps the EDM up to date and continuously valuable. Knowing the big picture saves design time and allows for each new application to fit together cleanly with existing applications. The CLDM proves to be an indispensible role in creating this big picture.

Conclusion

The Teradata CLDM saves organizations substantial amounts of time and money by providing a detailed and well-proven data model as a foundation for an organization's enterprise data model. In addition, the Teradata CLDM can be easily extended as the business grows, and provides the organization with a common understanding of business terms.

About the Author

Steve Hoberman is a world-recognized innovator and thought leader in the field of data modeling. He has worked as a business intelligence and data management practitioner and trainer since 1990, and is a popular presenter at industry conferences, both nationally and internationally. Steve is a columnist and frequent contributor to industry publications, as well as the author of *Data Modeler's Workbench* and *Data Modeling Made Simple*. He is the founder of the Design Challenges group and inventor of the Data Model Scorecard™. He can be reached at **me@stevehoberman.com**.

Data Model Scorecard is a trademark of Steve Hoberman & Associates, LLC. ERwin® Data Modeler is a registered trademark of CA, Inc. or one of its subsidiaries. SAP and R/3 are registered trademarks of SAP AG in Germany and in several other countries. Teradata is a registered trademark of Teradata Corporation and/or its affiliates in the U.S. and worldwide.