The Teradata Healthcare Industry Data Model Overview and Application

8.15 EB5890 HEALTHCARE

TERADATA.

Table of Contents

2 Introduction

4 Teradata Healthcare Data Model Overview

7 Enhancing the Healthcare Data Model with Unification

8 Teradata Healthcare Data Model Scenario

9 The Approach

15 Conclusion

15 About the Author

Introduction

As a small business owner, I needed to purchase health insurance for my family and me. Shortly after beginning coverage, my youngest daughter went for a medical checkup. Soon afterwards, we received a letter from the doctor's office stating our claim was denied because there was no record of us having insurance through our healthcare insurer. I called our insurer, and they confirmed we were not members. Checking through my financial statements however, I learned that all of our premium checks to our insurer had been cashed. I realized that this insurer had data integration issues between their financial management and claim areas, and therefore lacked a big picture view of their organization. This insurer lost any credibility in my mind, and I quickly switched insurance to another insurer, which, to date, appears to have a much stronger grasp on its data.

Knowing the big picture is critical to healthcare organizations for many reasons, including patient safety, profitability, and customer satisfaction. Every healthcare organization I've worked with has the noble (yet extremely challenging) goal of creating a high-quality electronic medical record and obtaining efficiencies within practice management. Getting the right information to the right people at the right time needs to be a reality, instead of just a saying. It can be achieved by having a single well-understood big picture of the organization. A single representation of member 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 a globe. Likewise, a data model can represent a specific functional area, such as supply chain, 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 that the concepts on a data model fit together as the CEO sees the company, as opposed to how individual functional or department heads view the company. There is one person who can play many roles including possibly being both a Provider and a Member. 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 lives in ten applications within an organization, the integrated EDM would show Person Last Name only once, and in addition, capture the mapping back to these ten applications, such as the person's last name as a Member, as a Provider, and as an Employee.

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 that the concepts on a data model fit together as the CEO sees the company, as opposed to how individual functional or department heads view the company.

There are resource and skill challenges with creating and maintaining an EDM, and therefore organizations are increasingly purchasing template 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 these Teradata Industry Data Models (iDMs):

- Teradata Communications Data Model (CDM)
- Teradata Financial Services Data Model (FSDM)
- Teradata Healthcare Data Model (HCDM)
- Teradata Manufacturing Data Model (MFGDM)
- Teradata Media and Entertainment Data Model (MEDM)

- Teradata Retail Data Model (RDM)
- Teradata Transportation and Logistics Data Model (TLDM)
- Teradata Travel and Hospitality Data Model (THDM)
- Teradata Utilities Data Model (UDM)
- Teradata Life Science Data Model (LSDM)

In the Teradata white paper titled, Leveraging the Industry Data Model, I provided an overview to the Enterprise Data Model and the Teradata iDMs. In this white paper, I will go into detail about the Teradata Healthcare Data Model (HCDM). Specifically, this paper provides an overview to the Teradata HCDM along with unification, and a scenario that illustrates how the HCDM can be leveraged. The goal of this paper is to increase your awareness of how the HCDM helps organizations obtain their big picture quicker 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 Healthcare Data Model Overview

The Teradata HCDM captures how a general healthcare organization works. It provides the big picture for a healthcare organization, containing more than ten broad subject areas, such as Claim, Campaign, and Clinical. I've studied industry models that were extremely generic, and therefore, only contained a handful of generic entities, such as Party. These generic models appear elegant yet require extremely complex mappings to the real source system to produce any value. The HCDM 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., an Incident Event, such as a Flu epidemic, that resulted in Professional and Pharmacy claims). These generalized constructs can even link across iDMs (e.g., Event and Party appear in other Teradata iDMs). Due to the details provided in the HCDM, the source system mapping becomes more manageable. The current version of the HCDM is extremely robust, containing more than 2,400 entities and 10,000 attributes, but these numbers—and model features—are continuously updated through new releases.

The HCDM is a living, breathing view of the healthcare business. This model provides a holistic view of healthcare insurers, providers, managed care organizations, healthcare data administrators, vendors, and consultants. 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 HCDM release. In July 2014, for example, Release 6.0 included integration of the ANSI X12/ASC 5010 industry standard concepts. Every Teradata HCDM iteration is a result of HCDM customers benefiting from the enhancement suggestions from many other HCDM implementations.

The HCDM 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 metadata. In addition, the HCDM documentation includes both hard copy and PDF files spanning four books. These include the Unified Data Models Reference Guide (we'll discuss unification shortly), a reference guide specific to the HCDM, a Physical Design Concepts Reference Guide, and two volumes of Appendices, which include support materials for the modeler; such as entity and attribute definitions, business data scenarios, abbreviations, and code tables. There are actually three ERwin files delivered with the product. These are a subject area model, a conceptual model, and the complete HCDM logical/physical data model. Each model provides an increasing level of detail, starting with the subject area model in Figure 1, which contains the high-level overview of the major subject areas covered.

The Teradata HCDM has a number of very important characteristics:

Operational

The HCDM captures how a healthcare company works instead of how a healthcare company typically does reporting. In other words, the vast majority of the structures in the HCDM capture the data elements and business rules that govern the day-to-day operation of the business. For example, the HCDM captures the business data where a Clinical Encounter (i.e., Patient Visit) may have led to a hospital Admission. In addition, there are sections of the HCDM that have been added to support analytics. Using this same Encounter and Admission example, there is a subject area for analytical models (e.g., risk scores) that can lead to a better understanding of Member risks for a particular disease, which could, in turn, lead to communications by the healthcare organization for proactive preventive healthcare reminders.

Logical / Physical

A logical/physical data model is a business solution for a specific set of business requirements. If a requirement is to capture claim information, the logical/physical data model would contain the data elements and business rules around the claim. 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

Teradata Healthcare Subject Area Model

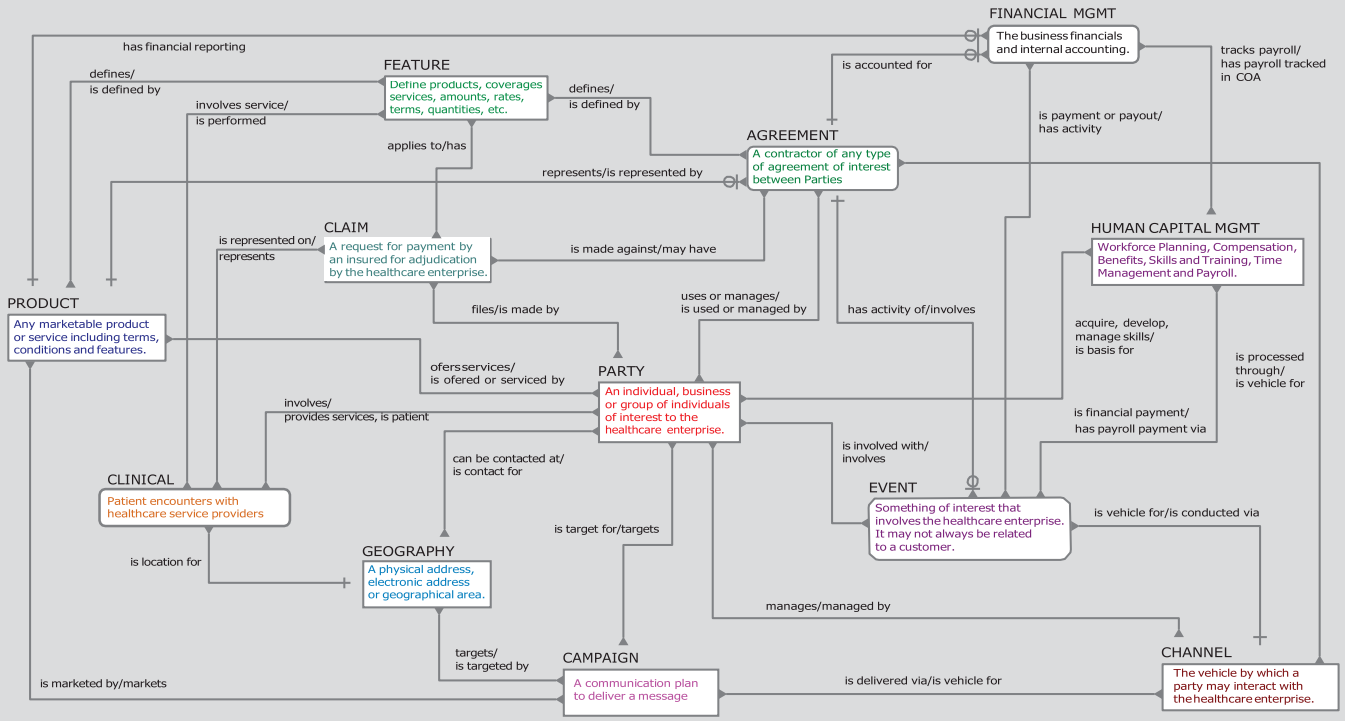

Figure 1.

primary key ("Every data element depends upon the key, the whole key, and nothing but the key.").

Extensible

The HCDM 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 HCDM as a starter model, and add new structures, defer any unneeded existing structures, and enrich the provided definitions to make them more meaningful to the organization.

Abstract

The HCDM 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. The HCDM can easily accommodate a new type of Event for example, as well as connect with other iDMs that also use the Event concept. This allows for greater commonality within and across the iDMs. All industries have Events for example, whether they are campaign solicitations in the banking industry, incidents in the healthcare industry, or service disruptions in the communications industry.

Global

The structures and terms on the HCDM 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'. The model contains data structures for the conversion of currencies and units of measure. It also allows for people and organizations to have many identifiers and sources of identification, such as country-specific tax identifications, passports, and drivers licenses. This facilitates communication on global projects and mappings back to global source systems such as ERP systems including SAP® R/3.

Standard

The HCDM consists of data elements in third normal form that support a number of industry standards, including Health Level 7 (HL7), American National Standards Institute (ANSI), Accredited Standards Committee (ASC) X12N, Centers for Medicare and Medicaid Services (CMS), Health Insurance Portability and Accountability Act (HIPAA), Health Plan Employer Data and Information Set (HEDIS), Joint Commission on Accreditation of Healthcare Organizations (JCAHO), Sarbanes-Oxley (SOX), American Society for Testing and Materials (ASTM), Continuity of Care Record (CCR), Joint HL7/ASTM Continuity of Care Document (CCD), and the National Uniform Billing Committee (NUBC). The data elements also follow best practice naming standards, including the use of class words based on the International Standards Organization (ISO) 11179 metadata 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 HCDM is sectioned into subject areas. Subjects are neatly captured in separate logical views (also called facets), and the use of color, distinguishing each subject area, makes it easier to digest the larger models. In addition, there are certain subject areas that are common across the iDMs, such as Party, Financial Management, and Geography. These subject areas have a common core in each iDM, and then are extended where appropriate within each of the models.

Enhancing the Healthcare Data Model with Unification

In today's economy where there is a tight relationship between creativity and survival, many healthcare organizations need to know more than healthcare. For example, healthcare organizations need a mastery of insurance and finance. However, the more we expand beyond healthcare, the greater the customization needed in the Data Model. However, by leveraging the similarities across all of the industry data models and creating building blocks that allow healthcare organizations to add on the relevant common component building blocks to create their unique models, the Teradata HCDM can better meet a healthcare organization's unique business model. This is the approach Teradata has taken with their unification project.

Unification is the integration of all iDMs creating a 'plug and play' industry data model environment. In addition to having separate iDMs for each industry, each iDM offers building blocks to the other iDMs. Roughly 45% of current iDM models are candidates for unification, and therefore, make up the building blocks, and about

55% remain specific to an industry. An organization still selects one of the industry-specific Data Models to represent their primary business, such as the HCDM. Along with this iDM comes a reference guide containing all of these building blocks that act as tightly coupled modules that allow for a high degree of sharing across the iDMs."

There are a number of benefits of unification including:

- More frequent model releases. Instead of receiving model releases annually, releases can be sent whenever a building block is enhanced.
- Less customization. By adding the necessary building blocks to an existing iDM, there is less model customization needed to fit a particular implementation.
- Greater applicability. As organizations continue to morph across industries, the iDMs can accommodate this.

So as a company's value chain increases, the iDM building blocks can be bolted on to provide support to cross those intersections.

Teradata Healthcare Data Model Scenario

HEU, a medium-sized health plan in the Midwest, has been consistently losing market share over the past five years. The CFO of HEU 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 they focus on premium pricing, reducing utilization costs, introducing new products, or retaining profitable members?

HEU has grown rapidly through acquisition of smaller health plans, and data integration has always taken a back seat. Many operational and reporting system silos 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 that 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 HEU'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 HEU's enterprise data management team, is responsible for building HEU'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 HCDM. The enterprise CDM was built leveraging Teradata's CDM that accompanies the HCDM

Types of Healthcare Data Models

Model	Purpose	All at once or incremental
Enterprise CDM	Capture the current understanding of the business on one piece of paper.	All at once
Enterprise CDM	Capture 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 2. Types of Enterprise Data Models

and contains the key entities (physical tables) within the HCDM PDM along with their relationships. The enterprise Data Model was built using the HCDM in four different roles, which we discuss later in this paper. The enterprise PDM was built based completely on the enterprise LDM/PDM. Figure 2 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 view of the organization using white boards 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 while standing at white boards and flipcharts. It represented each business area in their terms.

Three Different Definitions of Incident

An incident is any occurrence that causes a claim to be filed with HEU, examples being auto accidents, infectious disease outbreak (e.g., Lyme, West Nile Virus), or a Flu epidemic. [from the claims department manager]

An incident is any interaction between HEU and a person or organization outside of HEU, such as with patients, providers, or government agencies. [from the marketing department manager]

An incident is a financial transaction in which HEU is either the recipient or the payee of funds. Incidents include premium payments (HEU is the recipient) and remittance (HEU is the payee of funds). [from an accounting department representative]

Figure 3

The concept of Incident will be used to illustrate the four different types of models in this section. Incident is just one of the 2,400 entities in the HCDM (albeit an important entity), and it exists in the Event Subject Area. There were three different definitions of Incident identified in the white board CDM, as shown in Figure 3.

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 HCDM comes with a CDM that contains about 150 key concepts and their relationships for the healthcare industry. It was built by including several major entities from each subject area and then generalizing the rules among these remaining entities. For example, more than ten entities, including Claim, Health Claim, Institutional Claim, and Dental Claim, are represented by the single Claim entity on the Teradata CDM. Figure 4 contains a subset of the Teradata CDM.

The Teradata CDM allows an organization to achieve a high-level big picture of the organization without getting

Subset of Teradata Conceptual Data Model

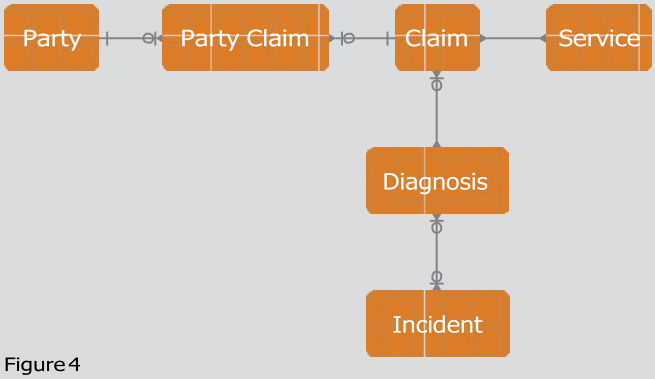

Figure 4

overwhelmed by jumping straight into a complex logical design. Jamie took a first pass at fitting the white board CDM into the Teradata CDM.

After spending time speculating how the pieces might fit together, Jamie organized a second series of meetings. These meetings took place in groups of ten to 15 individuals, and Jamie purposely invited people with very different views on 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 was either agreement on the model or issues that needed to be reconciled.

Figure 5 contains the portion of the Enterprise CDM after the terminology and definitions surrounding the term 'incident' were resolved.

The semi-circle with an 'X' in the middle is a subtype symbol. It identifies a grouping entity (in this case EVENT) called a supertype, as well as those entities sharing common data elements and relationships (in this case INCIDENT and FINANCIAL EVENT) called the subtypes. These three entities were already available in the Teradata CDM, each containing robust definitions.

The definition for EVENT in the Teradata CDM is: The EVENT construct tracks many types of interactions with the healthcare enterprise, such as financial (i.e., premium payments, remittance), communications (i.e., contact), HIPAA transaction (i.e., regulatory compliance), channel, incident (i.e., accident resulting in healthcare), etc. These events may involve patient care or be non-patient care related and may or may not affect a patient account. Events are related to Claims and Clinical Services, which allows tracking of a healthcare enterprise's

Subset of Enterprise Conceptual Data Model

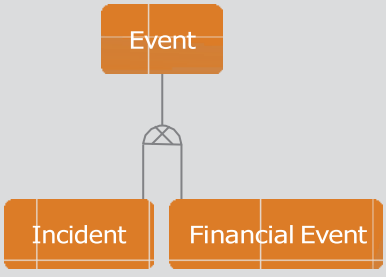

Figure 5

TERADATA.

financial position resulting from a patient encounter or an interaction with a party of interest with which the healthcare enterprise wishes to keep information.

The definition for INCIDENT in the Teradata CDM is: This is a subtype entity to EVENT and describes incidents that are of interest to the healthcare enterprise. This applies to events that cause (usually multiple) insurance claims, such as traffic accidents; natural disasters, such as hurricanes; and health episodes, such as a Flu epidemic. A major incident might also be referred to as a catastrophe.

The definition for FINANCIAL EVENT in the Teradata CDM is: This is a subtype entity of EVENT. This entity contains all events involved with the healthcare enterprise that are of a healthcare nature regardless of whether or not an account is involved. It includes maintenance transactions such as patient encounters. The involved customer is either the healthcare enterprise's account holder or a customer who does not have an account with the healthcare enterprise (a foreign customer).

The marketing department manager's definition for incident (recall Figure 3) most closely matched the definition for EVENT, and he was okay using the term EVENT; the claims department manager's definition most closely matched the definition of INCIDENT; and the accounting department representative's definition most closely matched the concept of FINANCIAL EVENT. In each of these cases, the definitions were expanded to include examples specific to HEU.

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

Enterprise Logical Data Model

As you might expect, the Enterprise 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 HEU's Enterprise LDM contained more than 700 entities and 1,900 data elements. It was built using a hybrid approach. A hybrid approach means it was built from both the top down and bottom up perspectives. 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.

Healthcare Data Model 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 HCDM (e.g., blueprint requires the most reliance on the HCDM and invisible the least). The degree of reliance is determined

Data Element Sample Mapping

Source			HCDM	
Source System	Table of File		Enity	
XYZ	Claim	Effective_Date	Event	Event Start Date
ABC	Marketing Plan	Begin_	Event	Event Start Date
X3000	Finance Plan	Start_	Event	Event Start Date

Figure 6. Types of Enterprise Data Models

by available modeling resources and knowledge of a particular business process.

Blueprint—the industry data model is *the* model. The HCDM 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: Healthcare organizations may model customer risk. Could also have customer scoring and segmentation models that describe the propensity of a customer to engage in a particular activity." In terms of HCDM, we can think of risk as potential healthcare expenditures for our members. By reducing risk through a proactive information environment, we can effectively reduce costs while improving our members' wellness. This is a win-win situation for both healthcare organizations and consumers. Analytical Model is a concept that the organization has not even considered relevant, yet after understanding its potential value of 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 HCDM including the actual Analytical Model entity.

Template—the industry data model is an integration point. The HCDM concept of Incident becomes an important integration point for the company. Each of the Incident data elements from the source systems was mapped into HCDM data elements. A sample mapping appears in Figure 6.

Note that this mapping is overly simplified, as usually there can be complex transformation rules, as well as other types of metadata that need to be reconciled, such as format, granularity, and nullability. This mapping does illustrate the usefulness of subtyping, as the Incident Start Date is really an Event Start Date (recall that an Incident is a subtype of Event).

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

Encyclopedia—the industry data model is referenced where needed. There is a need within the organization to understand product Features better and relate these features to the actual Product and a Claim. The HCDM provides a comprehensive data model for Feature and also includes in detail how Feature connects with Product and Claim. Jamie researched this area in the HCDM and was able to understand the data and rules behind the model so that she can add these concepts to the existing EDM. In some cases, terms, rules, and definitions from the HCDM 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 iDMs can serve as a reference to save some modeling time and reduce risk by ensuring all concepts are present on the model."

> 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 HCDM proves to have an indispensable role in creating this big picture.

Invisible – the industry data model is not consulted. The HCDM is not used at all. The Geography area is extremely well modeled within HEU and has been rigorously maintained for the past five years. For this area, the HCDM was not consulted at all. Parts of the HCDM that were used and contained geographic information were connected to HEU's existing party 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 HCDM came with a pre-defined PDM and more than one hundred typical business questions, and she used this list as a brainstorming technique with the business to agree on a set of common questions. In fact, one business question from the HCDM list became the scope for an entire data mart: "What are the per member per month (PMPM) analytics for total revenue, healthcare costs, and administrative costs?"

Smooth Sailing Ahead

All of HEU'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 HCDM proves to have an indispensable role in creating this big picture.

Conclusion

The Teradata HCDM 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 HCDM can be easily extended as the business grows especially by leveraging the Teradata iDM unification, and provides the organization with a common understanding of business terms.

About the Author

Steve Hoberman is one of the world's most well-known data modeling gurus. He understands the human side of data modeling and has evangelized "next-generation" techniques. Steve taught his first data modeling class in 1992 and has educated more than 10,000 people about data modeling and business intelligence techniques since then. Steve is known for his entertaining, interactive teaching and lecture style (watch out for flying candy!) and is a popular, frequent 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 Modeling Made Simple, Data Modeler's Workbench, and Data Modeling for the Business. He is the founder of the Design Challenges group and inventor of the Data Model Scorecard.® He can be reached at me@stevehoberman.com.

10000 Innovation Drive, Dayton, OH 45342 Teradata.com

8.15 EB5890

TERADATA.