Teradata Financial Services Data Model

Overview and Application

Table of Contents

1 Executive Summary

1 Introduction

2 Teradata Financial Services Data Model (FSDM) Overview

4 Enhancing the FSDM with Unification

5 Teradata FSDM Scenario

8 Conclusion

8 About the Author

Executive Summary

In the Teradata white paper titled, *Leveraging the Industry Data Model*, I provided an overview to the Enterprise Data Model (EDM) and the Teradata industry data models (iDMs). In this white paper, I focus on the Teradata® Financial Services Data Model (FSDM). This paper will provide an overview of the Teradata FSDM along with the unification framework and a scenario illustrating how to leverage the FSDM. The goal of this paper is to increase your awareness of how the FSDM helps organizations like yours obtain the big picture more quickly and accurately than building an EDM from scratch, thus permitting your organization to answer complex strategic and tactical business questions faster and more accurately.

Introduction

I worked four years as a data modeler and data architect for a Wall Street firm in its enterprise architecture group. As the data modeler, my role was to accurately capture requirements from analysis through design for a particular project. As the data architect, all of my project deliverables had to also be consistent with the enterprise. For example, if the project requirements needed "Customer," yet the enterprise as a whole used the term "Client," we would need to call it "Client" on the project design.

I found it very challenging to manage this enterprise perspective. Our group had a great manager who placed a high value on ensuring consistency between each project and the enterprise. However, many project managers that we worked with did not want to invest the additional time and cost associated with ensuring an enterprise fit. The end result of often-heated discussions was that the project would proceed inconsistent with the enterprise view.

During my tenure at this company, we purchased another financial services company, and there was an extremely time-consuming, painful, and costly effort to integrate systems. Our company was eventually purchased by a large global financial services company, and the integration issues were even more time-consuming, painful, and costly. Having a solid enterprise view would have definitely saved a substantial amount of time and money during both of these integration processes.

Knowing its big picture is critical for a financial services company. It can speed up analysis and design on specific projects, better comply with regulations and standards, such as Sarbanes–Oxley, Basel III, and Solvency II; reduce the effort and cost associated with mergers and acquisitions; and discover amazing things through business intelligence (BI).

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 risk management or it can represent the entire organization in the form of an EDM.

TERADATA®

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 currently offers these Teradata industry data models:

- Teradata Communications Data Model (CDM)
- Teradata Financial Services Data Model (FSDM)
- Teradata Healthcare Data Model (HCDM)
- Teradata Health and Human Services Data Model (HHS-LDM)
- Teradata Life Sciences Data Model (LSDM)
- 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 Financial Services Data Model Overview

The Teradata FSDM captures how a general financial institution works, covering retail and commercial banking, brokerage, investment, charge card, and property & casualty and life insurance business operations. It provides the big picture for a financial services organization, containing more than 300 broad subject areas, such as customer management, risk management, claims management, operations management, and finance and performance management. The FSDM is a living and breathing view of the financial services business.

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 FSDM release. The current version of the FSDM is extremely robust, containing more than 3,200 entities; 13,200 attributes; and 6,500 relationships.

The FSDM exists in an ERwin® Data Modeler file. ERwin Data Modeler is one of the more popular data modeling tools that offers reports for viewing and printing the models and their metadata. In addition, the FSDM documentation includes PDF files of several books. These include the Unified Data Models Framework Reference Guide (unification will be discussed later in paper), the Teradata Industry Data Model Physical Design Concepts Reference Guide, a reference guide specific to the FSDM, and two volumes of Appendices that include support materials for the modeler, such as a listing of the FSDM attributes, along with definitions, as well as entity and subject area definitions. The Reference Guide also contains a number of other sections, including a robust set of financial business questions.

The FSDM has a number of very important characteristics.

The "Top 10" reasons for making the move to the FSDM include (in no particular order):

1. **Provides agile teams with a solid foundation.** Out of all of the industries in which I have consulted, I have found that the financial services section relies most heavily on "fail fast" agile approaches. Agile approaches such as Scrum and Lean are practiced with the goal of producing working software in a minimal amount of time. Often however, short-term gains come with the price of long-term support. Agile teams at times produce useful project-oriented software that may not follow standards, good design principles, and enterprise architecture. Starting with the FSDM can dramatically reduce the amount of time to build a homegrown, consistent, and well-structured enterprise model to support agile teams.

2. **Links with Teradata Data Integration Roadmap**. Some organizations obtain the FSDM simply because of its connection with the Data Integration Roadmap. The questions that can be answered "out of the box" with the Data Integration Roadmap are important industry-level questions. Here are just a few of the hundreds of questions:

Customer Oriented:

- Which customers have a checking account but not a credit card account?
- Which customers have a homeowners insurance policy but not an automobile insurance policy?
- In which geographic area are my insurance customers concentrated?
- Which business segments, based on SIC codes, are growing? Diminishing?

- What type of activities, such as check deposits, are transacted mostly at a branch versus an ATM?
- What are the trigger points (such as a change in interest rate) that cause a loan to be paid off early?

Fraud Questions:

- What is the average deposit amount for a customer over time, and how does that compare to current deposit amounts? (Used to detect money laundering, for example.)
- Is the customer transacting at a location that is out of the normal geographic range? (Used to detect a stolen ATM card, for example.)

Account/Policy Questions:

- Which accounts have been securitized?
- What is the average application credit score for approvals and declines?
- How many insurance applications were accepted from a particular risk class?
- Which policies have customers not purchased and which they are likely to purchase?
- What is the number of accounts in a reward program, and points rewarded and spent?
- Which reward programs are most successful?

Risk Management Questions:

- How much capital should I allocate for credit risk purposes?
- What is the probability of default for a given account?
- What is the exposure of my internal investments to various European countries?
- What is the risk level and loss ratio by insurance agent?
- What are the insurance salvage, subrogation, and reinsurance recoveries on paid losses?
- What is the difference between the estimated loss at time of claim and the actual amount paid?

3. **Captures business processes and therefore is integration-friendly**. The FSDM captures an operational view instead of a purely reporting view. The operational view represents how an organization works as opposed to how an organization does reporting. Because an organization follows standard business processes, any applications that automate these processes should map to the FSDM with much less effort than if a reporting perspective is modeled. In addition, the model takes advantage of generic structures, such as Party and Event, to ease the mapping effort. On the FSDM, generic concepts like Party and Event are groupings for more granular and concrete parts of the business. For example, Associates and Households are two of the many more specific variations of a Party, and Payment and Claim are two of the many more specific variations of an Event.
4. **Minimizes risk**. When a data model is built from scratch, we sometimes question whether or not we have everything that is needed somewhere on the model. What would be the impact if a concept is accidently left off the model or represented improperly? Costs could include money, litigation, and/or business credibility. The FSDM is a result of more than 25 years of designing, implementing, and maintaining data warehouses for financial institutions, and is currently used in more than 175 financial institutions, a true testimony that it provides a trusted view.
5. **Facilitates information sharing**. If your business needs to exchange data with other companies within the same industry, the FSDM can be very valuable to establish a common language and set of rules. If two organizations need to share data, and they use the same terms and rules, it will make sharing much easier and less error-prone.
6. **Reduces maintenance costs**. When a new industry concept is introduced, Teradata includes this new concept on the next release of the model and therefore there is no maintenance cost to model the new concept from scratch. For example, the FSDM was enhanced to support the banking Basel III regulations for risk management. Teradata experts took the Basel III Accord regulations and identified the business data elements paragraph-by-paragraph. These business data elements were then mapped to the current FSDM to identify the gaps, and the FSDM was modified to include these new regulations. Another example: The FSDM was modified significantly in the collateral area with respect to the valuation and adjustments to the value of the collateral, the relationship between collateral items and collateral agreements, and the relationship between the collateral agreements to the loan (exposure) agreements. The model had to account for the fact that a piece of collateral could be pledged against many loans within the bank and also outside the bank. Teradata experts have done the same kind of analysis for Solvency II (insurance) regulation. Another area of modification was securitization–loans that are packaged and then offered as a financial instrument to investors.

Securitization by banks removes the risk from the bank's books. The FSDM can now trace the financial instrument back to the loan package (group of loans) and back to the individual loan. Not having this securitization tracking process across the business chain was one reason for the economic crisis. The underlying loans of many of these packages were high risk, but the trail was lost between the bank, the special purpose vehicle (financial entity) that packaged the loan, and the investor. The FSDM was modified to handle securitization agreements from the perspective of the selling bank and the investing bank. The FSDM now has the concept of "tranches," which is the piece of the group of loans offered as a financial instrument.

7. **Gracefully extends to accommodate new concepts.** The FSDM 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 FSDM can easily accommodate a new type of Event, for example, as well as connect with other Teradata 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, emails to change a shipment date in the manufacturing industry, or service disruptions in the communication industry. Abstraction has been applied in not just industry-neutral terms such as Event, but also for terms specific for finance and insurance. For example, instead of specifying all of the dates associated with a claim, the generic concept of Claim Date Type allows for all types of dates to be easily accommodated such as Occurred Date, Settled Date, Receipt Date, and any new types of dates, as well.
8. **Provides global and standard perspective**. The structures on the iDMs are designed for international use, and not just United States-based. For example, the term "postal code" is chosen over "ZIP code." In addition, the data elements on each iDM follow best practice naming standards, including the use of class words based on the 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. So the class word for Customer Last Name is "name."
9. **Presents the concepts in digestible chunks**. The FSDM is sectioned into subject areas. Subjects are neatly captured in separate views, and the use of color distinguishing each subject area makes it easier to digest the larger models.
10. **Contains unification**. Unification allows organizations the option to select the iDM that most closely represents their business and then "bolt on" pieces from other iDMs to create a more representative picture of an organization. (Read more about unification in the next section.)

Enhancing the FSDM with Unification

Rarely can traditional labels define organizations accurately. That is particularly true in the current economy as many branch out to perform functions that cross traditional industry and market boundaries. More and more, companies simply do not fit any specific mold:

- Retailers may have a credit card line.
- Car manufacturers may have a finance arm for loans.
- Oil and gas refiners may do oil and gas trading (commodity trading).
- Utilities may sell electricity (also commodity trading).

Therefore, even when they know the benefits, diversified companies often say they cannot find a single suitable industry model that works best for them. Fortunately, it is now possible to build a unified data model to see beyond the enterprise view to the holistic value of an organization.

Companies typically choose an iDM that is tailored for their industry. But to be at all useful, the model needs to accommodate changes such as expansion into different markets. That might require two or more iDMs that must be either integrated or heavily customized. A better method is to select one primary iDM and then add the relevant components from other industries to create a unique model. For example, a utility company may choose the Teradata Utilities Data Model (UDM) as their primary iDM, yet leverage the modules of Account, Event, Channel, Party, Geography, and Forecasting, to handle the commodity trading side of their business. This method of unification captures the best practices of industries and creates a plug-and-play environment that employs

elements of the separate iDMs as building blocks for others. Consequently, the models reflect real-world applications while providing a single view of the operation to support growing business plans.

In general, roughly 45 percent of the content of a Teradata industry model is a candidate for unification with the other industry models. By leveraging these building blocks, an organization can select one industry Data Model to represent its primary business. That one industry model then becomes the foundational reference guide or framework with which the other building blocks or independent modules are integrated.

The benefits of a unified data model include:

- **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, organizations need less model customization to fit a particular implementation.
- **Greater applicability**. As organizations continue to morph across industries, iDMs can adapt and accommodate their needs.

Teradata FSDM Scenario

A medium-sized insurance company in the Midwest primarily writes three lines of business (LOB): Auto, Residential, and small commercial or Business Owners Policies (BOP). Each business unit currently operates autonomously, and senior management has as a core business objective to integrate analytics and reporting for all three lines with the goals of more efficient operating processes, and more powerful and useful analytics information. Today, even relatively simple reporting questions, such as "How many of our auto insurance policyholders are self-employed and therefore potential candidates for business liability insurance?" can only be answered after much manual effort. Often there is a high degree of error in the results. Last year, for example, this insurance company conducted a direct mailing campaign to personal lines customers whom they believed were good prospective business liability insurance policyholders, offering them an attractive multi-policy discount. However, because of a high degree of errors in answering cross-LOB questions, this insurance company mistakenly offered the multi-policy discount to more than 200 existing business liability insurance policy holders who did not qualify for the offer.

Another opportunity area for this insurance company is to understand the exposure and risk associated with natural disasters such as earthquakes, floods, or hurricanes. Although today there is some analysis done in the residential insurance area, such as additional premiums required if a house is near an earthquake fault line or in a flood zone, this insurance company does not have the big picture and cannot answer, for example, "What is my exposure across all lines of business if there is a Class 5 hurricane that will hit the southeast?"

To answer broad business questions such as these requires an integrated enterprise data model within the scope of an enterprise architecture. Although there are knowledgeable business and information technology staff within each LOB, there are no resources that have expertise in all three business areas. Senior management therefore decides to fund a Center of Practice (COP) containing a senior business analyst and data architect from each LOB to work together to create an enterprise architecture.

The three business analysts and three data architects meet and realize even the most basic concepts vary in definition and sometimes terminology. For example, returning to the natural disasters question earlier, each line of business classifies the reason for the claim differently. Because of this it is impossible to determine broad exposure risks.

For example, in the automobile line of business, there is only a general classification used, which is an indicator capturing whether the claim was a result of human negligence or a result of Mother Nature. The business division does have a more robust classification scheme than the automobile LOB, yet it is done based upon characteristics rather than events. So for example, if a hurricane causes damage to a small business that has its business policy with this insurance company, instead of capturing the hurricane event, this insurance company captures the characteristics of the incident such as the claim was a result of wind or rain. The residential unit does have the most robust way of capturing the cause of the claim, and they use Claim Reason Codes for this purpose. For example, a Claim Reason Code 55 is for a hurricane and a Claim Reason Code 25 is for a broken water pipe. So right away even with this first example, there are major integration issues. How can we answer the previously-raised question, "What is my exposure across all lines of business if there is a Class 5 hurricane that will hit the southeast?" if each line of business categorizes the cause of claims differently?

Senior management knew that some more standardized business processes would be coming out of this six-person group, and would therefore support changing existing business processes. The six-person team decides to therefore use the FSDM as an integration point for their operational systems and as a foundation for the integrated data layer within their data warehouse program. So they started with the FSDM term of Incident. A subset of the definition for this term from the FSDM is:

This is a subtype entity to EVENT and describes incidents that are of interest to the financial institution. 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 heart attack. A major incident might also be referred to as a catastrophe.

The group decided terms like Incident should be used organization-wide, and incidents are events, and events tie to people, organizations, locations, policies, and time periods. Terms such as these are therefore ideal for integration.

Since integration is an activity that needs to involve different levels in the organization, from business professionals through the very technical roles of database administrators and developers, there needs to be different views of the data as captured through different levels of data models. The group decided therefore that four data models were needed: an enterprise subject area model, an enterprise conceptual data model, and both logical and physical data models of the data warehouse-integrated data layer. Together the group crafted explanations and purposes for each of these models, as shown in Figure 1.

Enterprise Subject Area Model

The FSDM comes with a Subject Area Model that contains ten key concepts and their relationships for the financial services industry. See Figure 2. It captures the concepts and relationships across the FSDM's broad subject areas. 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

Model	Purpose	Built All at Once or Incrementally
Enterprise Subject Area Model	A communication tool for the Top 10 concepts and their business rules within the organization. Useful for explaining core concepts and for explaining the similarities across the different business units.	All at Once
Enterprise Conceptual Data Model	Although more detailed than the Enterprise Subject Area Model, this model still fits on a single piece of paper, and contains more industry-specific terms, yet at a high enough level where the units fit together holistically. Useful for communicating scope of large initiatives such as the data warehouse program.	All at Once
Integrated Data Layer Logical/Physical Data Model	A very detailed and normalized view of the business, not tainted with technology such as software and hardware constraints. Useful for ensuring the data warehouse contains the information to answer business questions, both questions within an LOB as well as cross-LOB questions.	Incrementally, driven by project requirements
Integrated Data Layer Physical Data Model	The integrated data layer PDM tainted with technology, such as software and hardware constraints. Useful for working with database administrators and developers to ensure the business questions are being answered with the most efficient structures and process. Along with this model, we also need to have a spreadsheet containing the mapping from source application data models to this model.	Incrementally, driven by project requirements

Figure 1. Types of Data Models Needed.

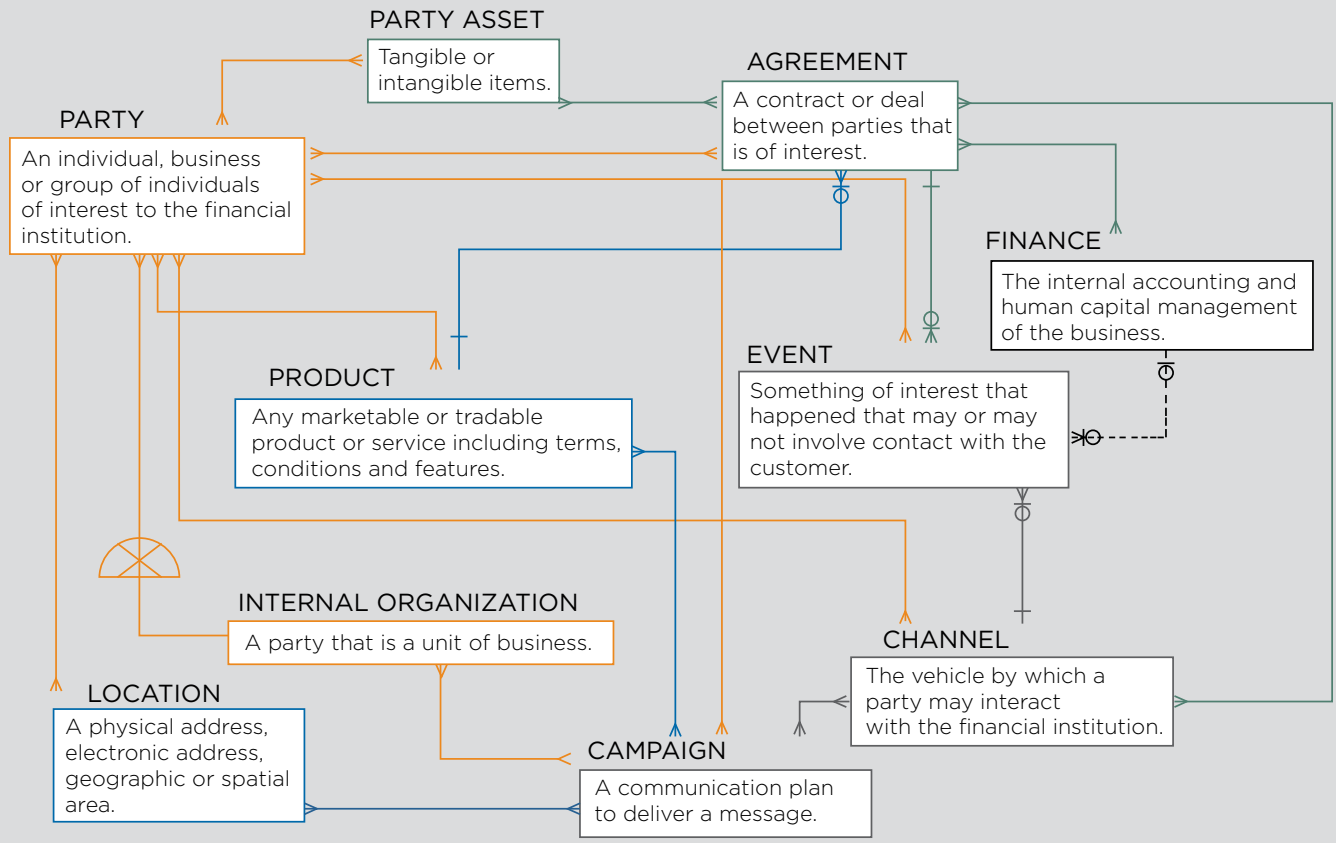

Figure 2. Teradata FSDM Subject Area Model.

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 easy traceability. For example, all of the Party entities in the FSDM are red in the Subject Area Model, and also in the Conceptual and FSDM models.

It is typical to customize industry data models based on need, and therefore even on this high-level data model some of the names and definitions need to be modified for this particular insurance company. For example, this insurance company prefers the term "Sales Channel" over "Channel," and so this entity was renamed on this model.

Enterprise Conceptual Data Model

The conceptual data model contains more detail than the subject area but yet still at a high enough level to be easily readable on a single piece of paper. The FSDM also comes with a conceptual data model, and this model contains about 200 entities, as opposed to the ten on the subject area. Much of the additional detail on the conceptual comes through subtyping, a small example shown for the Event concept in Figure 3.

Integrated Data Layer Logical Data Model

The integrated data layer logical data model for this example scenario 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. Many business analysts found it challenging to extract questions from the business folks. Luckily, the FSDM and the Data Integration Roadmap came with several hundred typical business questions, which were used as a brain-storming technique with the business to agree on a set of common questions. As you might expect, the integrated data layer LDM required more effort than the prior two models. It had more detail and required the most discussions to resolve the integration issues.

The FSDM was the starting point for each project, and the project teams often extended the model themselves with the Architecture Center of Practice group reviewing any modeling work to ensure consistency.

Figure 3. Subset of FSDM Conceptual.

Integrated Data Layer Physical Data Model

The Enterprise Physical Data Model (PDM) was also built incrementally on a project-by-project basis, driven from the logical data model. A combination of views, indexing, and de-normalization were frequently-deployed techniques to ensure that the reporting was user friendly with quick response times.

Smooth Sailing Ahead

All of the insurance company's future operational and BI applications relied on the Enterprise Conceptual Data Model (CDM) as a starting point for design. When each application data model was considered complete, a review took place to identify possible Enterprise CDM changes as a result of this application model. So each application starts with the Enterprise CDM and then contributes new ideas back to the Enterprise CDM. This keeps the Enterprise CDM 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 FSDM proved to have an indispensable role in creating this big picture.

Conclusion

The Teradata FSDM saves organizations substantial amounts of time and money in building a data warehouse by providing a detailed and well-proven data model as a foundation for an organization's enterprise architecture. In addition, the FSDM can be easily extended as the business grows (especially by leveraging the unification framework) and provides the organization with a common understanding of business terms.

About the Author

Steve Hoberman is the most requested data-modeling instructor in the world. In his consulting and teaching, he focuses on templates, tools, and guidelines to reap the benefits of data modeling with minimal investment. He 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, spanning every continent except Africa and Antarctica. Steve is the author of five books on data modeling, including the bestseller *Data Modeling Made Simple*. He is the founder of the Design Challenges group, inventor of the Data Model Scorecard,® and CEO of Technics Publications. Steve can be reached at me@stevehoberman.com, @DataMdlRockStar on Twitter, or through Steve Hoberman on LinkedIn.

10000 Innovation Drive, Dayton, OH 45342 **Teradata.com**

04.15 EB6546