

Industry logical data models

Are they right for your enterprise? *by Steve Hoberman*

Before you use a tool, you should understand how it functions. For business intelligence (BI), this principle means that before you can use data for reporting and analysis and for making timely business decisions, you need to understand how the business—and the industry—works.

The right tools to capture how an organization within a particular industry works, or should work, are essential for in-depth knowledge of your business. These tools should also provide expert modeling structures for your enterprise to use as a guide. An industry logical data model (iLDM) offers this kind of information and guidance.

Defining iLDM

Mapping and integrating a company's data is not a new thing: Enterprise data models (EDMs), which contain all of the data produced and consumed across an entire organization, have been around for more than 20 years. Until recently, however, these models, sometimes called "corporate models" or "business models," had to be built by each organization from the ground up.

The iLDM is a pre-built model that extends the EDM concept. Like the EDM, it is a subject-oriented and integrated view of the information needed for an enterprise data warehouse (EDW) to answer strategic and tactical business questions. Rather than stop with the information within an organization, however, the iLDM reaches across the

industry, offering a commonality among organizations within that industry. Companies not able or willing to build their own EDM can utilize this pre-built model and customize it to fit their specific organization.

As a logical data model (LDM), an iLDM is application-independent. This means software, hardware and communications constraints are not present. While it does not contain history, the iLDM represents a snapshot of information (point in time). It also does not contain derived data or other calculations to assist with retrieving data more quickly. All model changes necessary for issues like speed, storage, security, backup and recovery must be added at the physical level. This is beneficial because by not imposing

organizations requires most companies to customize about 25% of the model, which includes adding data elements and enriching definitions. Only the key data elements appear on the models; therefore, any other data elements must be added to the appropriate entity. In many cases the starter definitions provided will require additional text to be meaningful to a specific company.

To allow for additional flexibility and commonality across iLDMs, a fair amount of abstraction (combining things under generic terms, such as Event and Geography) is contained within each model. All industries have Events, for example, whether they are bookings in the travel industry or orders in the manufacturing industry.

Like the EDM, the iLDM is a subject-oriented and integrated view of the information needed for an EDW to answer strategic and tactical business questions.

hardware, software or reporting constraints directly in the model, each organization can first understand its business and then modify this business view to perform efficiently within its unique technical environment.

Common information that companies share within an industry are contained within the iLDM. A more complete solution for

For viewing ease, each iLDM is sectioned into individually colored subject areas; even the conceptual data model (CDM) is color coded by subject area. Certain shared subject areas (SSAs) are common across iLDMs. Extended where appropriate within each of the models are SSAs such as Party, Event, Geography and Demographics, which have a

common core in each iLDM.

The iLDM structure is designed for international use; for example, the term “postal code” is chosen over “ZIP code.” The data elements on each iLDM follow best-practice naming standards, including the use of class words based on the ISO 11179 metadata standard. Class words, such as Name, Code, Identifier, Date, Quantity and Amount, represent the high-level category of a data element. So the class word for the data element Customer Last Name is Name.

For a simplified example of an iLDM, see the figure below, which displays six entities common across the iLDMs that belong to the Geography SSA.

In this example, an Address can be either a Telephone Address, Electronic Address or Mailing Address. There are certain data elements that are unique to each of these subtypes, such as the Telephone Area Cd (Code), which belongs only to a Telephone Address. A Dwelling Type can belong to zero, one or more Mailing Addresses, and a Mailing Address can belong to zero or one Dwelling Type.

Provided within an iLDM are certain definitions, which may be customized to make them meaningful in your organization. Below are examples of some standard definitions and a possible customized substitute.

Standard:

- > Dwelling Type entity: Identifies the type of dwelling. Example: single family, multi-family and so on.
- > Dwelling Type Cd data element: A code that uniquely identifies the DWELLING TYPE.

Customized:

- > Extended definition of Dwelling Type Cd: A code that uniquely identifies the DWELLING TYPE, which is a place of residence or a Post Office Box. Place of residences include single-family, multifamily and temporary housing.

When to acquire an iLDM

Industry logical data models are relatively new and not yet considered a common practice or a necessity among industries, but organizations that have employed iLDMs are experiencing a

competitive advantage. It is important to understand the factors that can sway an organization to use an industry model and the factors that can convince an organization to build its own EDM.

An organization might use an iLDM for these reasons:

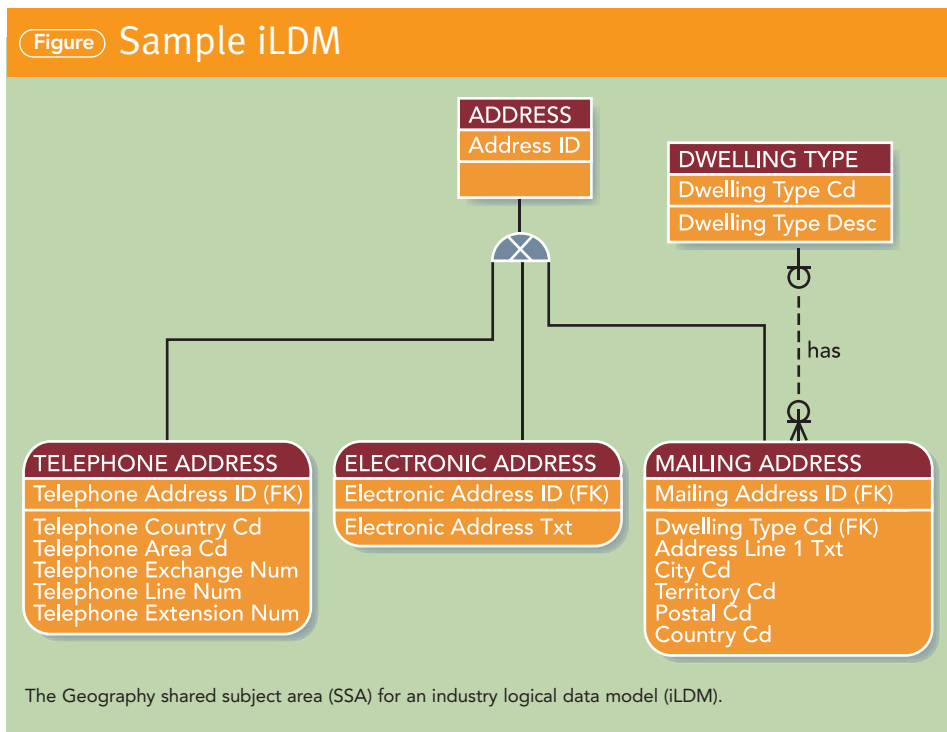
- > **Start-up considerations like time and money.** If your organization needs to show results quickly, employing an industry model can give you a head start. Rather than designing a model starting from scratch, you can map your existing structures to the model, then customize it to fit your business.

> **System maintenance.** Many data-modeling efforts fail to consider the routine maintenance required of an EDM. Since data constantly changes, the EDM will quickly become obsolete if it is not regularly kept fresh and accurate. An outside resource, on the other hand, helps keep the industry model up to date for you—you only need to maintain organization-specific changes. For example, when radio frequency identification (RFID) was first developed, the Teradata Manufacturing iLDM (MLDM) updated its model, which organizations were able to use without maintaining it themselves.

> **Risk and modeling expertise.** Building an EDM requires skilled resources from both business and IT. An industry model lets you leverage someone else’s expertise.

> **Existing application infrastructure and vendor relations.** Some vendors offer industry models that work seamlessly with their other products. Teradata iLDMs, for example, work with the Teradata EDW Roadmap, which creates a strong connection between the EDM and BI solution.

> **Information sharing within your industry.** Data is commonly shared among multiple companies in an industry—an excellent reason to leverage an industry model. Since organizations in the same industry (such as the rental car companies and airlines in the travel industry) apply



Industry logical data models are relatively new and not yet considered a common practice or a necessity among industries, but organizations that have employed iLDMs are experiencing a competitive advantage.

similar constructs and terminology, using an iLDM will allow for the same terminology across organizations.

Below are four reasons why organizations might build their own LDMs:

> **Limited requirements.** If your EDM only needs to exist at a subject-area level, you may not need an industry model. Subject-area models take a relatively short amount of time to create and maintain, thereby maximizing the time- and money-saving benefits.

> **Existing EDM.** If you already have an EDM, or have a good start to one, it might be easier to use that rather than a pre-built one at this stage. (Note: If you have differing views regarding how to name or to define concepts, a pre-built industry model provides an unbiased external view.)

> **No suitable industry models exist.** There are certain niche industries, such as energy, where no industry model exists. Customizing a pre-built model to fit your business needs would be overly challenging, time consuming and costly, so it would be more efficient to create a new model.

> **Pioneering culture.** If your company believes in building everything itself, the EDM is usually no exception.

Factors listed in both of the previous sections impact whether an organization decides to acquire an industry model or create its own in-house. For example, a manufacturing organization, Company X, already used an EDM to drive an integrated data warehouse; plus it had modeling expertise—both strong points against

purchasing an industry model. The company's pre-existing EDM was also rich in attribution and definitions, another valid reason to stay with its own system. But the EDM was a collection of physical data modeling structures (not logical)—which is difficult to reuse and integrate.

Because of the need for an integrated logical view, and since it had a relationship with Teradata as a customer, Company X obtained the Teradata MLDM. It currently uses the MLDM as a model template and enhances this template with definitions and additional attributes from the existing pool of physical tables.

This gives the company the best of both worlds: The MLDM provides the integrated flexible view, and the existing models provide the richness to the definitions, additional data elements and business rules. Today, more than 100 entities in its data warehouse are based on the MLDM.

Company Y, however, was given two weeks to build an enterprise view. There was no budget allocated to purchase an industry model and no budget allocated for maintaining the completed EDM. Due to time and money constraints, Company Y decided to build a subject-area EDM. Each subject area on the model was mapped back to the source system responsible for creating that data. Each project team that needed to build a new application was required to reference this subject-area EDM as a starting point for its more detailed logical analysis. In the long run, this proved more costly, more time consuming and less productive for its system.

The best bet for an industry model

The most useful industry models are those that are detailed, accurate and heavily attributed. Vendor modelers who build these iLDMs are typically highly experienced within a given industry.

As a global leader in enterprise data warehousing, Teradata consults with many companies, which yields experience that leads to a continually deeper and broader understanding of many industries. Teradata Professional Services has documented this business knowledge in the form of seven distinct iLDMs. These specific industries are communications, financial services, healthcare, manufacturing, retail, transportation and travel.

The Teradata iLDMs use proven modeling methodologies and an experienced professional services team to avoid the common and expensive pitfalls that can arise with building an EDM from scratch. Using this expertise in an iLDM enables organizations to focus on their business and keep ahead of their competition. **T**

Steve Hoberman has worked as a BI and data management practitioner and trainer since 1990 and is 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. You may visit his Web site at www.stevhoberman.com.



Visit Teradata.com to download Steve's white paper, "Leveraging the Industry Logical Data Model As Your Enterprise Data Model."