



Enterprise Data Modeling: 7 Mistakes You Can't Afford to Make

How ER/Studio® Can Help You Avoid Them

Karen Lopez, InfoAdvisors



InfoAdvisors

© InfoAdvisors • All Rights Reserved

Karen Lopez
@datachick

www.infoadvisors.com

Helping You Add Business Value to Your Information Management Resources

#TEAMDATA

An updated version of *Enterprise Data Management: 7 Mistakes You Can't Afford to Make*

This white paper is sponsored by IDERA Inc.
ER/Studio is a registered trademark of Embarcadero Technologies, Inc., a
wholly-owned subsidiary of IDERA Inc.

Executive Summary

Today's information technology professionals are faced with a variety of challenges including legislation, organizational change, and the rate of technological innovation. In our experience most IT professionals in organizations initiating an enterprise data modeling program feel overwhelmed at the scope and complexity of initiating new methods, tools, and techniques. They are eager to deliver the benefits of well-managed enterprise architecture but have concerns about possible missteps along the way. With the advent of Big Data, NoSQL and new data technologies adding complexities to an already confusing data stack, it's no wonder data professionals are feeling the pressure to squeeze even more value out of data resources.

Whether their organization is just starting out or experienced in enterprise data modeling efforts, there are certain pitfalls that can become obstacles to success.

Before we talk about the pitfalls, though, we should discuss terminology.

Enterprise Data Modeling

encourages a practical balance between enterprise and project points of view

comprises both application and enterprise data models

enables IT groups to respond more quickly and effectively to business needs

delivers information that is the most useful to the business

uses the proper tools and techniques in delivering project outcomes

Enterprise data modeling includes the use of logical and physical data models as a key component of enterprise data architecture. These models form a knowledge base of data and information requirements in the context of the business as a whole.

This paper first looks at the benefits of an effective enterprise data modeling effort. Why, if there are pitfalls, should an organization approach data modeling with an enterprise point of view? Let's start first with the reason information technologies exist—to enable businesses to accomplish goals that would be nearly impossible with manual processes. Enterprise architectures allow teams to improve communications, reduce and better align information technology expenditures, and develop better project estimates and plans. However, the most

ENTERPRISE DATA MODELING: 7 MISTAKES

important benefit of an effective enterprise data modeling program is that it instills a greater confidence in the information technology group's ability to deliver solutions to the business.

The second section of this paper discusses seven common mistakes that organizations can make in developing enterprise data models. These mistakes each have a cost that negatively impacts projects and the information technology group as a whole. For each potential mistake, we show how IDERA Inc.'s ER/Studio Enterprise Team Edition can ensure your teams are able to deliver results to the business.

Benefits of Enterprise Data Modeling

Why do organizations develop enterprise data architectures?

Most data management professionals have experience modeling at a project level. Their success is evident in project-specific data models used for developing new databases. However, some organizations have difficulties leveraging these activities to deliver enterprise-wide benefits of a more integrated effort. In order to do so, they need to ensure they have an effective enterprise data management process.

Enterprise data management is the management of data and information in a manner that supports a business's goals and objectives.

By *enterprise data management*, we mean the management of information in a manner that encourages a practical balance between enterprise and project points of view, contributes to an overall enterprise architecture, delivers information that is the most useful to the business and uses the proper tools and techniques to deliver successful project outcomes. It includes using logical and physical data models as a component of the architecture. These models form a knowledge base of data and information requirements in the context of the business as a whole. Ensuring that projects produce enterprise-class models is the key to a successful enterprise data management program.

An enterprise data architecture is an organized model of a business's data and information assets. Enterprise data management involves the production of formalized enterprise data architecture. Like all architectures, it may have a graphical representation of its contents, as well as descriptive information about data and information. This descriptive information is *metadata*. An enterprise data architecture specifies the properties, rules, and policies for managing an organization's data.

An enterprise data architecture specifies the properties, rules, and policies for managing an organization's data.

Let's look first at the benefits and outcomes of a successful enterprise data management effort.

Improved Communication

Managing enterprise metadata increases the value of business information by serving as a *lingua franca*, a standardized language, easily understood by everyone from business users to application developers to database administrators.

Data comes from a wide variety of business systems, which in turn are based on a wide variety of technologies. The product of all this variety is a Tower of Babel that prevents business users from leveraging their information.

A user, looking for information about product profitability, can't find the quarterly sales figures for some of his company's new services. He discovers that since the new services are being managed in a purchased application package, all the intelligence on the new services is hidden away in separate databases and spreadsheets. He spends several days getting access to the new data sources and trying to understand their structures and still does not have a reliable method for reporting the quarterly sales figures.

Another user has distributed a scathing report to an organization's executive committee and various vendor partners showing that the partners are missing targets for service levels and have been missing them for quite a while. As a series of urgent meetings are conducted, the original user learns that the data item he used to calculate service levels, Completed Time, reflected the date and time the work was entered into a system, not the time the work was actually finished. The user is embarrassed, the business partners are frustrated, and the executive committee has lost faith in the ability of their employees to properly report progress.

With effective enterprise data management, the users just described would have had a single reference point for all information artifacts, along with descriptions and definitions, giving them better access to their own data.

Reduced and Better Aligned Costs

In an environment where projects evolve independently of an overall enterprise architecture, a great deal of time is spent by team members attempting to discover sources of project data, understanding its meaning, and confirming its accuracy and timeliness. Individual team members may develop an understanding of the enterprise's data, but time constraints and lack of adequate tools mean that there is little opportunity to share that data across project teams. Even if documentation is developed, it can be difficult to consolidate the knowledge contained therein unless it was developed with a goal of reuse—which is unlikely.

However, for a team working in an organization managing data at the enterprise level there are more opportunities to reuse previously developed models, potentially saving weeks or months of effort.

A project team is in the early stages of developing an application to support warehouse functions and surveys the enterprise model for objects relevant to their project. While enterprise modeling is relatively new to their organization, they are able to “check out” approximately 35 percent of the project's modeling objects from a previous team's efforts—employee, location, organization, accounting, inventory, product and vendor information. Since previous projects had taken an enterprise point of view in developing their models, a significant portion of the current project team's work has already been completed.

Another project team working on the acquisition of a software package to support the organization's invoice processing function will be able to make use of modeling objects from both projects. While this team may not be building a database, it will be able to make use of the models as requirements for evaluating packages as well as inputs to package evaluation criteria. Each object in a model becomes a requirement to be met by candidate packages.

The fact that these project teams are working in an environment that encourages an enterprise approach not only shortens time spent on analysis and design efforts but also leads to reduced costs. Team members can focus on the quality of their deliverables due to the time saved in reusing models.

As projects use and supply models to the enterprise architecture, they are contributing to a repository of objects that can be used when estimating future projects. During the estimation phase of a project, modelers can contribute schedule and resource estimates based on their inventory of modeling objects. If the subject to be addressed has a good deal of coverage from previous projects, then the effort to complete the data models for that project will be less than that of a project that has had very little coverage.

Our modeler from the development project in the previous section would review a project charter's scope section to determine the extent to which it encompasses existing modeling artifacts. He could then estimate resource and time needs based on the existence of prior work. By using tools to see what concepts have been implemented, what concepts have been designed but not implemented, or what teams are using similar concepts, he can more accurately plan for both design and implementation resources.

Working from a clear and consistently developed enterprise architecture means project teams can apply clear and consistent estimates and plans.

Increased Confidence in IT Efforts

The most important benefit of an effective enterprise data management effort is increased confidence in the IT resources of an organization. Saving costs and offering quality models are part of that confidence, but they alone do not advance the overall belief that an organization's IT department is doing what is right for the business.

Better Partnerships with Subject Matter Experts

Often business users find that every time they initiate a new project or a refinement to an existing project, they are forced to retrain the IT staff as requirements they have contributed in the past were not properly recorded.

Without the benefits of a formal data management program, a significant portion of a knowledge worker's or a business analyst's time is spent trying to understand data and how it

can be used. Users know that the business participated in the design of the systems and they don't understand why the IT department has no idea what data exists, where it is used, and why it is so hard to find.

Several business users are assigned to a project that is developing a new inventory tracking system that will provide a key competitive advantage. A team of IT professionals is assigned to the project to determine the requirements, turn them into designs, and eventually develop a new system. Several weeks of full-day workshops in addition to cumbersome review cycles are required. Eventually, after many false starts and misunderstood requirements, the system goes into production.

Fast forward a year—the kinks are worked out of the system and operations are running smoothly. A change in legislation, a new line of products and a new organizational structure lead to a new project to enhance the system. Many of the IT staff have been reassigned to other projects, moved on to new opportunities, or have retired. The business users are frustrated to find out that they have been scheduled, again, for a series of workshops to determine what they do, what policies must be followed, what exceptions there are, and what data constraints are required to ensure the quality of the data. Many users fail to show up for the workshops because they believe the time they spent a year ago was wasted; they have little motivation to participate in the new project.

What if the project team were able to provide a complete set of the business requirements previously captured, review them with the users who have the changes in mind, and have revised models available in a fraction of the time that would have been required if starting from a blank piece of paper? Wouldn't that leave the business users with a feeling of confidence for future resource requests?

Faster and More Accurate Response to Change

Another influence on the business's perception is the IT organization's ability to respond to legislative or regulatory changes.

Newer legislative efforts and public confidence with regard to information privacy and antiterrorism efforts have had a direct impact on data management activities. Organizations must now explain how data will be protected, guarded, and disclosed. Many laws specifically identify information attributes that require special management.

For example, many data breach laws include disclosure notification requirements for social security, bankcard, and driver's license numbers. These attributes are found in many information systems, even in the systems of businesses that do not have core functions of social programs, banking, or licensing. Privacy legislation also focuses on specific personally-identifiable data. As required in these laws, organizations must properly manage sensitive data and track who uses it and how it is used.

ENTERPRISE DATA MODELING: 7 MISTAKES

A disgruntled employee walks out the door one night with two production servers. During the investigation into the missing hardware and the information stored there, management asks, “exactly what data resides on those servers?” and “are we compliant with the data protection clauses of disclosure of compromised information?”

A data architect reviews the physical models for the databases and other data stores for the missing servers. She determines that while some sensitive data was included, all of it was encrypted as required. She provides a list of the kinds of data stored in the databases so that the customer service and public relations organizations, with the guidance of the legal department, can handle inquiries.

An enterprise data architecture provides a robust framework for supporting business inquiries about the location and status of data.

Increased Professional Respect

In addition to having a strong process, it is important to have a flexible and feature-filled data management tool. Business users and IT professionals want to have models that are readable, customizable, flexible, and above all, understandable.

In our experience, models that are well presented and walkthroughs that are well run greatly increase confidence in IT. Using an enterprise data management approach to IT activities provides team members with a knowledge base of the corporation – a resource that can be used for planning, evaluating, and developing compliance activities. It increases business’s confidence that its projects will be well run, its time will be well used, and the results will be completed at a professional level.

The Pitfalls

How they happen and how to avoid them...

Now that we've covered the benefits of effective enterprise data management, let's look at the pitfalls that must be avoided. Even experienced IT professionals make mistakes, but using the proper tools and techniques, we can ensure that we don't make the most serious ones. In discussing each mistake, we also show how ER/Studio Enterprise Team Edition from IDERA Inc. can help avoid those mistakes.



Mistake 1: Forgetting that an enterprise architecture is a living framework

Some IT professionals think of data architecture as a final, fixed deliverable instead of a versioned view of the business. Many project plans arrive with a start and end date for the data modeling effort, where the logical model ends on a Thursday and the physical model starts (and can be completed) on Friday, with no tasks or resources assigned for refinements. This sort of finish-to-start mentality can lead to the perception of incomplete tasks and projects, and unmet business needs, as refinements are a natural and expected part of any modeling effort. This type of pitfall is almost guaranteed if non-data management professionals develop project plans in isolation.

If team members do not understand that data models, like other project deliverables, can be versioned, shared, and reused, they are likely to misunderstand the role of models during a project's lifecycle. Every phase of a project can lead to refining previous decisions, understandings, and changes due to external influences.

Any team member should be able to trace a business concept from the logical model to the physical model to the physical implementation of that concept. This traceability is a key to realizing the benefits of an enterprise data management program.

Many of the features of ER/Studio are designed to facilitate refinement and traceability:

Roundtrip Modeling

ER/Studio Data Architect allows modelers and designers to make changes to the logical model, generate changes to the physical model as needed, and apply the changes to databases. Using the *Where Used* feature allows anyone to see how a concept is modeled and implemented. ER/Studio also allows changes made in a database to be compared to the models. A modeler can then decide whether these changes should be reflected in the logical and physical models.

Compare and Merge

ER/Studio Data Architect can compare models to databases, other models, and DDL. The comparison results allow the modeler to decide how to address differences at a detailed level—apply to first model, apply to second model/database, or ignore. *Compare* and *Merge* functions provide modelers with the ability to share enterprise-class objects across projects.

Versioning

Models and objects can be versioned with comments when adding them to the repository. Models also can be checked into the repository as a snapshot, allowing for a complete point-in-time view of a model version.

Universal Naming

It is common in the midst of a modeling effort to refine a naming standard, or to decide to change all references for one word to another word. For example, an organization may decide that references to *CLIENT* should be changed to *CUSTOMER*. The *Universal Naming Utility* makes this an easy change to make. The utility finds all occurrences of *CLIENT* and on a case-by-case basis, or in bulk, changes it to *CUSTOMER*.

Denormalization Mapping

ER/Studio supports common physical model changes such as table splitting (horizontal and vertical), subtype rollups and roll downs, column duplication from one table to another, and table merging. All modifications completed through the denormalization mapping function retain their link to the logical model object, allowing modelers to follow the changes from logical to physical via the *Where Used* feature.

The reality is that models and their contents will need to change when the business changes, when technology constraints and limitations are modified, and when additional analysis has been completed. Appropriately responding to changes increases business confidence in IT organizations.

Appropriately responding to changes increases business confidence in IT organizations.

Mistake 2: Keeping data models invisible



An invisible model might as well not exist. If a data management effort is going to deliver business value, it must be accessible, understandable, and shareable. Models need to be available in an easily searchable manner, including definitions. A model without definitions is just a diagram that could be interpreted in many ways, so it is important to supply the entire model with definitions and appropriate metadata.

Models need to be available in an easily searchable manner, with both diagrams and metadata.

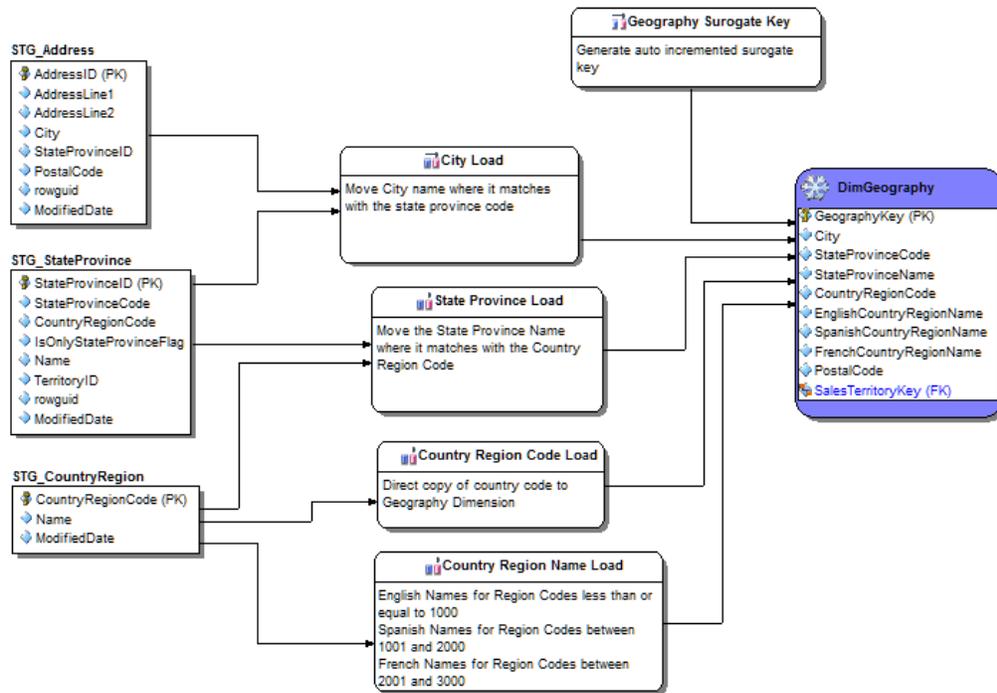
It makes sense that the more available models are, the more likely they will be used by a wide variety of IT and business professionals.

ER/Studio features can help a data management team ensure that models are always available:

Data Lineage

ER/Studio allows analysts to show how data from one system forms the input for another system, such as a data warehouse. The Data Lineage feature shows source, transformations and target data structures right within the data model.

ENTERPRISE DATA MODELING: 7 MISTAKES



Collaboration via Repository

With ER/Studio's Repository, teams can share their models while maintaining control over who performs updates, who can check them out, and who can check in changes. Administration of these rights can be done by project roles, thereby simplifying model management.

Enterprise Data Dictionary

ER/Studio supports, via the Repository, enterprise definitions of highly shareable data. For example, a modeler can define standard properties for *Person Name*. This definition would be managed within the enterprise data dictionary and shared with all models, ensuring that project teams can build models and systems that are integrated.

Team Server

ER/Studio Team Server allows IT and business users to know their data by giving them the tools to search data models and metadata via an intuitive web interface. This self-service application provides cross-model access via the Repository. End users and IT team members can contribute metadata and provide feedback on corporate glossaries, business terms, data models and their components.

Mistake 3: Assuming that business users can't understand or review models

3 This pitfall typically happens when a project team has a history of poor preparation for model walkthroughs. Business users will resist reviewing models if they do not have the proper training, or if presenters have unclear and inconsistent methods for presenting the models.

Most business users can understand and review models if they are given proper training, if they are guided through a structured review process, and if they have regular access to the models. It is common on our projects for users to insist that data models be used. They insist on these models due to training they've attended, and they understand the importance of data models. One or two days of user-targeted training can expose these valuable team members to the issues a modeler faces in creating and maintaining a model, as well as the goals and benefits of an enterprise approach to modeling.

ER/Studio Viewer

IDERA offers a read-only version of ER/Studio, which allows users to view and print models and reports. ER/Studio Viewer provides access to models without the risk of introducing inadvertent changes.

HTML Reporting

ER/Studio can produce an interactive, Web-based "intranet" that includes clickable diagrams, definitions, and some or all the metadata for each modeling object. This browser-accessed site provides access to anyone who needs to understand what data has been modeled and where it is used. The Web-based models can be configured to tailor the amount and presentation of metadata depending on the audience. For instance, business users might be interested in objects and definitions, while developers may need access to more detailed implementation properties.

Team Server

With Team Server's cross-model search facility, business users can find concepts even when they don't know the technical name of a data item. For instance, a user could search on the term *customer*, finding that word in a definition or other field of the table *CS_CLNT_CST*.

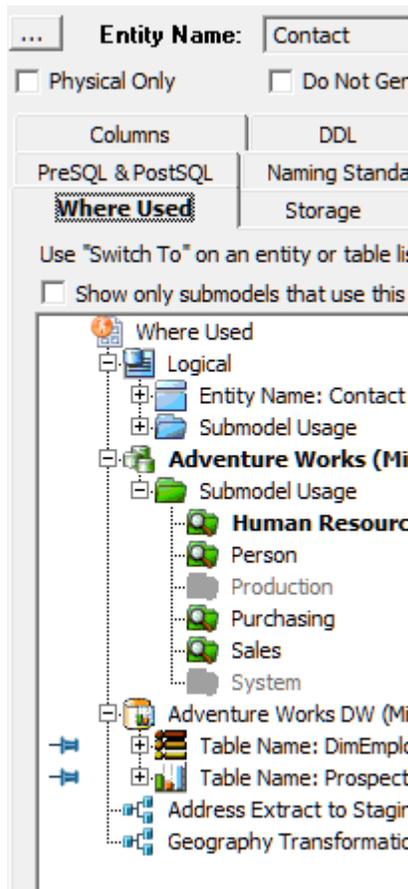
Business users who see and work hands-on with models regularly are more likely to support the allocation of resources to future efforts.

Mistake 4: Thinking that data models are only about databases



Both logical and physical data models can be used for more than supporting databases. Logical models, with their technology independence, organize data requirements in a manner that no text-based document alone could do. Physical models allow team members to see how business concepts are implemented without having to provide access to production systems.

Data models are not just a means to building a database. They can be used by a team performing package evaluations. They can be used to develop Requests for Proposals and evaluation criteria. They can be used as training materials for new staff or as references for *ad hoc* queries.



Where Used Analysis

Users that have access to *ad hoc* query tools can review a model’s “where used” metadata, together with definitions, to determine the best source of information for their queries.

Analysts can view the complete set of models and recruit objects for use on new projects. This understanding of these enterprise assets can be used in project planning and estimating.

Attachments

One of the most helpful features of ER/Studio is the ability for a modeler to extend models to include customized metadata. The Attachments feature allows a team to add metadata for each type of modeling object and to selectively apply values to each instance. For example, if a team wants to document security and privacy information, they might add an Attachment for privacy level codes to a submodel and a sensitivity level code to each entity.

Metadata Import and Export

Modelers can easily export valuable metadata via ER/Studio’s metadata export facility. This feature allows data to be saved in formats that can be opened in dozens

of different applications, including other modeling and database tools, XML schemas, desktop applications, and more. Metadata Import and Export ensures that data management professionals are able to work collaboratively with other architecture professionals no matter what tools they use.

Macros

With ER/Studio’s built-in *macro* features, modelers can easily develop functions that automate the extraction of metadata to desktop productivity applications such as word processing and

spreadsheet applications. This feature allows a team to develop customized reports, review scripts, progress reports—just about anything that can be used to present enterprise metadata.

Thinking only of databases means missed opportunities for sharing the value of data architecture components.

Mistake 5: Throwing models “over the wall”



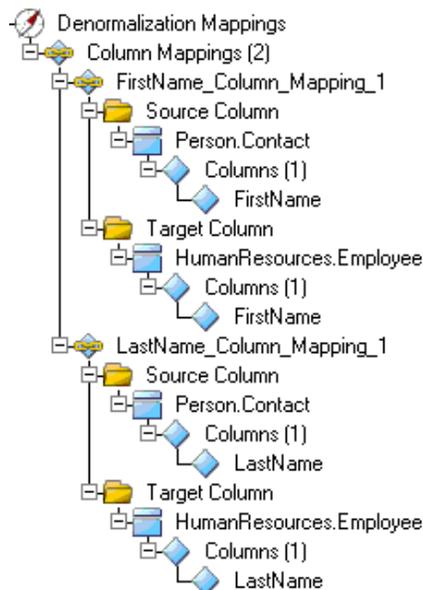
As we have seen previously, some IT professionals see data models as deliverables that are completed early in the project, and then handed off to others to decide how and when they use the models.

This approach is almost always guaranteed to cause a breakdown in communications between the business needs and what is implemented. A modeler should be involved in how the requirements are captured as well as how they are implemented. She should understand where they are implemented, and what changes, if any, have been made due to technical constraints or performance requirements. A data modeler is the mediator between business requirements and physical implementations.

A “throw it over the wall” approach will lessen the value of data models, as round-trip modeling is almost always going to be a manual process, prone to errors and slowing down the benefits inherent in enterprise data management.

Denormalization Mapping.

Denormalization is a natural part of physical design. ER/Studio’s ability to maintain mappings between the logical model and its potentially different physical structure can be used to ensure that physical design changes stay true to the original intent of the logical model.



Synchronized Models and Databases

ER/Studio’s ability to generate and modify databases ensures that there is a consistent link between the models and the physical implementations of the models. Using data model-driven development methods ensures that business requirements are accurately reflected in system implementations.



Mistake 6: Forgetting about the sizzle

Since one of the main benefits of effective enterprise data management is better communication, it is critical that presentations of models be clear and understandable. Model diagrams should include guides so the non-modeler can follow the intent and meaning of modeling objects. They should be easy to read and well commented. Models should be interesting. Never underestimate the value of sizzle.

Models should be interesting. Never underestimate the value of *sizzle*.

Use of Presentation Features

In many cases the use of color, including distinctive shades of grey, to highlight certain modeling objects can be effective in communicating metadata to a reviewer. For example, one could use color to indicate stewardship information, modified entities, or even priorities for review. Fonts, sizes, shading—all of these presentation features can sharpen the focus of data model reviews and uses.

It is even possible with the *macro* feature to automatically color modeling objects based on user selections, custom metadata (attachments), or even the completeness of modeling.

Use of Automation

The *macro* features in ER/Studio can help modelers extend and customize modeling efforts to meet the unique needs of their project or entire enterprise models.

```
ObjType = MySelObject.Type
If ObjType = 1 Then
    ' Get the ID for the selected object.
    ID = MySelObject.ID
    ' Get the actual entity object with this ID.
    ' The model contains the collection of all the
    ' entities.
    Set MyEntity = MyModel.Entities.Item(ID)
    If Logical = True Then
        ' If the model is logical:
        ' Get the entity name.
        EntityName = MyEntity.EntityName
        ' Set name of PK and PK attribute.
        AttributeName = EntityName + " ID"
        PrimaryKeyName = EntityName + " PK"
```

Use of Diagramming Objects

ER/Studio has features to add drawing objects to a model to highlight areas, call attention to objects, or add notes to the diagram. These drawing objects add quality to the models by facilitating user reviews, calling attention to modifications, or adding information to ensure the proper communication of business rules.



Mistake 7: Thinking of them as “your” models

Finally, the most critical mistake that can be made is treating data models as if the modeler personally owns them. The models should be presented as belonging to the business and tended to by the modelers. That means sharing them openly, providing access to those who want it, keeping extra printouts available, offering training on how to read them and making every effort to make them clear and understandable.

Models and their underlying metadata should be seen as corporate assets to be managed by a partnership of modelers and the business.

Models and their underlying metadata should be seen as corporate assets to be managed by a partnership of modelers and the business. If they are portrayed as just technical specifications that can be understood only by developers and database administrators, they will not provide the benefits of an enterprise data architecture.

The enterprise-class data modeling tool features described earlier can further support the business’s ownership of the models and their uses beyond databases—their presentation, their role in a living, thriving enterprise architecture, and the use of sizzle where needed—so they become a natural and beneficial part of delivering solutions to the business.

Finally...

We hope that you see the importance of an enterprise data management in promoting better data integration and project coordination while reducing project risks. Benefits such as enhanced communication and increased confidence in the IT organization are crucial to meeting business needs— especially needs imposed by external sources. The incremental benefits of reduced cost and faster response to change are a natural product of taking an enterprise approach to data management.

Costs and benefits aren’t the only things we need to manage. Risk is a natural part of all projects. Organizations that understand potential risks are better positioned to avoid them. InfoAdvisors and ER/Studio can play important roles in mitigating the pitfalls of a new enterprise data modeling effort.

About the Sponsor



IDERA understands that IT doesn't run on the network – it runs on the data and databases that power your business. That's why we design our products with the database as the nucleus of your IT universe.

Our database lifecycle management solutions allow database and IT professionals to design, monitor and manage data systems with complete confidence, whether in the cloud or on-premises.

ER/Studio is the collaborative data modeling solution for data professionals to map and manage data and metadata for multiple platforms in a business-driven enterprise data architecture.

Whatever your need, IDERA has a solution.

About the Author

InfoAdvisors

Karen López is Senior Project Manager and Architect at InfoAdvisors. She has more than twenty years of experience in helping organizations implement large, multi-project programs. She is a Microsoft MVP, SQL Server, and a Dunn and Bradstreet MVP.

InfoAdvisors is a Toronto-based data management consulting firm. We specialize in the practical application of data management. Our philosophy is based on assessing the cost, benefit, and risk of any technique to meet the specific needs of our client organizations.

InfoAdvisors offers data modeling training, including training focused on introductory data management issues, for non-modelers such as business users, database administrators, project managers and developers.

Find Karen at www.infoadvisors.com