CHAPTER 3
What camera settings also apply to a data model?

Cameras have settings
Zoom, Focus, Timer, Filter
Data models, too

This chapter compares the data model to a camera, exploring four settings on the camera that equate perfectly to the data model. Understanding the impact these settings can have on a data model will increase the chances for a successful application. This chapter also compares the camera's film to the three levels at which the data model can exist: subject area, logical, and physical.
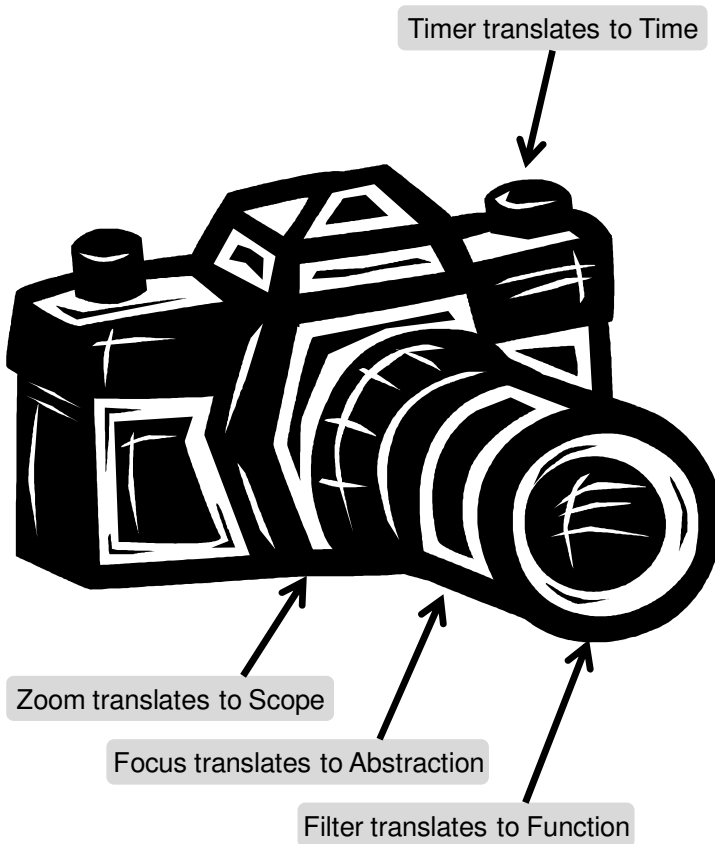
The Data Model and the Camera

A camera has many settings available to take the perfect picture. Imagine facing an awesome sunset with your camera. With the same exact sunset, you can capture a very different image based on the camera's settings, such as the focus, timer, and zoom. You might for example zoom out to capture as much of the sunset as possible, or zoom in and focus on people walking by with the sunset as a backdrop. It depends on what you want to capture in the photograph.

There are four settings on a camera that translate directly over to the data model: zoom, focus, timer, and filter. A model is characterized by one value from each setting. See Figure 3.1.

Figure 3.1 Camera settings that translate to data model variations

Timer translates to Time

Zoom translates to Scope

Focus translates to Abstraction

Filter translates to Function

The zoom setting on the camera allows the photographer to capture a broad area with minimal detail, or a narrow scope but with more detail. Similarly, the scope setting for the model varies how much you see in the picture. The focus setting on the camera can make certain objects appear sharp or blurry. Similarly, the abstraction setting for the model can use generic concepts such as Party and Event to "blur" the distinction between concepts. The timer allows for a real-time snapshot or a snapshot for some time in the future. Similarly, the time setting for the model can capture a current view or a "to be" view sometime in the future. The filter setting can adjust the appearance of the entire picture to produce a certain effect. Similarly, the function setting for the model adjusts the model with either a business or application view.

And don't forget that the type of film you use is important! A proof sheet shows all of the images on a single piece of paper, the negative has the raw format of the image, and the output can be in any one of a number of formats, including paper film, slide, or digital. Similarly, the same information image can exist at a subject area, logical, or physical level of detail on a data model.

Which setting is right for your model? As with photographing the sunset, it depends on what you want to capture. Match the goals of your model with the appropriate model settings to improve the overall quality of the data model and the application it supports.

Scope

Both a data model and photograph have boundaries. Boundaries determine what will be shown. A photograph can capture my youngest daughter enjoying ice cream (actually her whole face enjoying the ice cream), or the photograph can capture my daughter and her surroundings, such as the ice cream shop. Similarly, the data model can include just claims processing, or it can include all concepts in the insurance business. Typically, the scope of a data model is a department, organization or industry:

- **Department (Project)**. The most common type of modeling assignment has project-level scope. A project is a plan to complete a software development effort, often defined by a set of deliverables with due dates. Examples include a sales data mart, broker trading application, reservations system, and an enhancement to an existing application.
- **Organization (Program)**. A program is a large, centrally organized initiative that contains multiple projects. It has a start date and, if successful, no end date. Programs can be very complex and require long-term modeling assignments. Examples include a data warehouse, operational data store, and a customer relationship management system.
- **Industry.** An industry initiative is designed to capture everything in an industry, such as manufacturing or

banking. There is much work underway in many industries to share a common data model. Industries such as health care and telecommunications have consortiums where common data modeling structures are being developed. Having such a common structure makes it quicker to build applications and easier to share information across organizations within the same industry.

Abstraction

A photograph can be blurry or in focus. Similar to how the focus on a camera allows you to make the picture sharp or fuzzy, the abstraction setting for a model allows you to represent "sharp" (concrete) or "fuzzy" (generic) concepts.

Abstraction brings flexibility to your data models by redefining and combining some of the data elements, entities, and relationships within the model into more generic terms. Abstraction is the removal of details in such a way as to broaden applicability to a wider class of situations, while preserving the important properties and essential nature of concepts or subjects. By removing these details, we remove differences and, therefore, change the way we view these concepts or subjects, including seeing similarities that were not apparent or even existent before. For example, we may abstract Employee and Consumer into the more generic concept of Person. A Person can play many Roles, two of which are Employee and Consumer. The more abstract a data model, the fuzzier it becomes.

On a data model, concepts can be represented at different levels of abstraction: 'in the business clouds', 'in the database clouds', or 'on the ground':

- **In the business clouds**. At this level of abstraction, only generic *business* terms are used on the model. The business clouds model hides much of the real complexity within generic concepts such as Person, Transaction, and Document. In fact, both a candy company and insurance company can look very similar to each other using business

cloud concepts. If you lack business understanding or do not have access to business documentation and resources, a model 'in the business clouds' can work well.

- **In the database clouds**. At this level of abstraction, only generic *database (db)* terms are used across the model. The database clouds model is the easiest level to create, as the modeler is "hiding" all of the business complexity within database concepts such as Entity, Object, and Attribute. If you have no idea how the business works and you want to cover all situations for all types of industries, a model 'in the database clouds' can work well.
- **On the ground**. This model uses a minimal amount of business and database cloud entities, with a majority of the concepts representing concrete business terms such as Student, Course, and Instructor. This model takes the most time to create of the three varieties. It also can add the most value towards understanding the business and resolving data issues.

Time

Most cameras have a timer, allowing the photographer to run quickly and get in the picture. Similar to how the timer on a camera allows you to photograph a current or future scene, the time setting for a model allows you to represent a current or future "to be" view on a model.

A model can represent how a business works today or how a business might work sometime in the future:

- **Today**. A model with the today setting captures how the business works today. If there are archaic business rules, they will appear on this model, even if the business is planning on modifying them in the near future. In addition, if an organization is in the process of buying another company, selling a company, or changing lines of business, a today view would not show any of this. It would only capture an 'as is' view.

- **Tomorrow**. A model with the tomorrow setting can represent any time period in the future, and is usually an idealistic view. Whether end of the year, five years out, or 10 years out, a tomorrow setting represents where the organization wants to be. When a model needs to support an organization's vision or strategic view, a tomorrow setting is preferred. I worked on a university model that represented an end of year view, as that would be when a large application migration would be completed. Note that most organizations who need a tomorrow view have to first build a today view to create a starting point. But that's ok! Just as a photographer can take more than one picture of a scene, so, too, can the data modeler build more than one data model with different setting values.

Function

Filters are plastic or glass covers that, when placed over the camera lens, adjust the picture with the color of the filter, such as making the picture more bluish or greenish. Similar to how a filter on a camera can change the appearance of a scene, the function setting for a model allows you to represent either a business or functional view on the model.

Are we modeling the business' view of the world or the application's view of the world? Sometimes they can be the same and sometimes they may be very different:

- **Business**. This filter uses business terminology and rules. The model represents an application-independent view. It does not matter if the organization is using a filing cabinet to store its information, or the fastest software system out there; the information will be represented in business concepts.
- **Application**. This filter uses application terminology and rules. It is a view of the business through the eyes of an application. If the application uses the term 'Object' for the term 'Product', it will appear as 'Object' on the model and it

will be defined according to the way the application defines the term, not how the business defines it.

Format

A camera has a number of different formats in which the photo can be captured. The format setting adjusts the level of detail for a model, making the model either at a very broad and high level subject area view, or a more detailed logical or physical view:

- **Subject area**. Often when a roll of film is processed, a proof sheet containing small thumbnail images of each photograph is included. The viewer can get a bird's eye view of all of the photographs on a single sheet of photo paper. This bird's eye view is analogous to the subject area model (SAM). A SAM represents the business at a very high level. It is a very broad view containing only the basic and critical concepts for a given scope. Here, basic means that the subject area is usually mentioned a hundred times a day in normal conversation. Critical means that without this subject area, the department, company, or industry would be greatly changed. Some subject areas are common to all organizations, such as Customer, Product, and Employee. Other subject areas are very industry or department specific, such as Policy for the insurance industry or Trade for the brokerage industry.
- **Logical**. Before the days of digital cameras, a roll of processed film would be returned with a set of negatives. These negatives represented a perfect view of the picture taken. The negative corresponds to the logical data model. A logical data model (LDM) represents a detailed business solution. It is how the modeler captures the business requirements without complicating the model with implementation concerns such as software and hardware.
- **Physical**. Although a negative is a perfect view of what was taken through the camera, it is not very practical to use. You can't, for example, put a negative in a picture frame or in a photo album and easily share it with friends.

You need to convert or 'instantiate' the negative into a photograph or slide or digital image. Similarly, the logical data model usually needs to be modified to make it usable. Enter the physical data model (PDM), which is the 'incarnation' or 'instantiation' of the LDM, the same way as the photograph is the 'incarnation' of the negative. A PDM represents a detailed technology solution. It is optimized for a specific context (such as specific software or hardware). A physical data model is the logical data model modified with performance-enhancing techniques for the specific environment in which the data will be created, maintained, and accessed.

EXERCISE 3: Choosing the Right Setting

In the following table, check off the most appropriate settings for each of these scenarios. See the Appendix for my answers.

1. Explain to a team of developers how an existing contact management application works

Scope	Abstraction	Time	Function
❑ Dept	❑ Bus clouds	❑ Today	❑ Bus
❑ Org	❑ DB clouds	❑ Tomorrow	❑ App
❑ Industry	❑ On the ground		

2. Explain the key manufacturing concepts to a new hire

Scope	Abstraction	Time	Function
❑ Dept	❑ Bus clouds	❑ Today	❑ Bus
❑ Org	❑ DB clouds	❑ Tomorrow	❑ App
❑ Industry	❑ On the ground		

3. Capture the detailed requirements for a new sales datamart (A datamart is a repository of data that is designed to meet the needs of a specific set of users)

Scope	Abstraction	Time	Function
❑ Dept	❑ Bus clouds	❑ Today	❑ Bus
❑ Org	❑ DB clouds	❑ Tomorrow	❑ App
❑ Industry	❑ On the ground		

Key Points

- There are four settings on a camera that translate directly to the model: zoom, focus, timer, and filter. Zoom translates into data model scope, focus into the level of abstraction, timer into whether the data model is capturing an 'as is' or future view, and filter into whether the model is capturing a business or application perspective.

- Match the goals of your model with the appropriate model settings to improve the overall quality of the data model and resulting application.

- Don't forget the film options! Would your audience prefer to view the proof sheet (subject area model), the negative (logical data model), or the photograph (physical data model)?