3.3. Entity level

Overview

In the previous modules we discussed several issues related to representing an information object. That discussion assumed we knew what we meant by an information object. This module reviews the concept of entity level and its purpose in developing a representation scheme for a collection of information objects.

What is an entity?

In this course, we define an entity as any thing about which we can make assertions, where assertions can be in the form of statements about the attributes (properties or characteristics) of the entity. As an example, a person is an entity. We can make an assertion about the entity of person that it has the property of two arms, two legs, etc. We can also make descriptive statements about the entity such as "the person has blue eyes." A group of people can be an entity. We can make an assertion about the entity of the group that it has the property of regular meetings (e.g., it's a book club). We can make descriptive statements about this entity such as "all members read the same book each week."

One of the differences between the entity of the person and the entity of the group is that they represent different entity levels. We can make statements about one of the entities that wouldn't make sense to make about the other. For example, the group entity has the property of a regular meeting schedule, but the person entity does not.

Entity level

When we create a representation of an information object, we need to be clear as to the scope, extent, size, etc. of the object we are creating. We use the concept of entity level to address this. The terms unit of analysis and unit of description are conceptually equivalent to the term entity level.

We can define entity level as the chunk of recorded information about which we will be making assertions or statements. The representations we have discussed in previous modules comprise a set of assertions or statements we will make about an entity. Thus, we must have a clear sense of the entity level our representations describe.
For example, a library catalog record typically represents the entire book, not chapters or stories in a book. The entity level represented by the catalog record is a single book. The Reader's Guide to Periodical Literature has entries (i.e., records) that represent individual articles from magazines and journals. The entity level represented by a single entry in the Reader's Guide is an individual article.

While both these entity levels might have common attributes (e.g., they each have a title, an author, a subject), there may be some attributes that only apply to a book but not to an article (e.g., the book may have a table of contents and index, but it is unlikely the article would have these parts). Thus, the entity level may prescribe what statements we can make about an information object.

**Choosing an entity level**

In many cases identifying an entity level to which the representations will apply may be obvious. In other cases, you may want to carefully consider exactly the chunk of information you will represent in a single record.

Let's say you have a collection of music recordings on compact disks. There are at least two possible entity levels you could choose. One level is the compact disk as a whole. The other level is the individual songs on the disks in the collection. In the first case, the physical container of the compact disk becomes the defining feature of the entity level. Your representation would include information about the CD as a whole, with information about it and its contents. In the second case, the individual song (or works as we'll begin calling them) are the chunks that get represented. In one case you have a single record representing the whole CD, and in the other case you have a single record representing an individual song. Which is the right entity level? In this case, there is not a right or wrong choice about entity level but simply an awareness of the implications for searching and retrieving information. If you represent the whole CD, and users are searching for individual songs, how do the songs get represented when the entity level is the whole CD?

Another format that often comes up is videos. Often the physical video cassette contains just one movie (i.e., work). Yet sometimes, a physical cassette contains two or more discrete works (e.g., a video cassette that has two Beatrix Potter stories). One choice would be to represent the physical cassette, no matter how many works it contains. Another choice would be to represent the individual works. These are issues that the concept of entity level helps us to address.

**Project Alert!** In IOP section 2.1. you have to explicitly state the entity level of your organization system. This is to demonstrate that you are clearly aware of what a single database record represents. An information object when viewed from different entity levels may prescribe different attributes about which to record information. The choice of entity level may seem obvious at first, but it's important for you to realize the implications of your choice. It is good to include in your narrative an explicit statement that what you choose as an entity level is represented by one record in your database.

While we have become accustomed to some of the bibliographic tools for searching for information and have come to expect to see representations of certain entity levels (e.g., we expect in the library catalog to see representations of whole books), not all users necessarily understand our bibliographic tools. Pity the poor user who goes to the online catalog searching for a specific article in *Time* magazine; not understanding the entity level of a particular bibliographic tool (like the catalog) can leave the user in a quandary.
An organization system that has no explicit statement of the entity level represented can also cause problems for the catalogers and record creators. When a video cassette comes into the collection, is the record created by the cataloger representing a work on the video cassette or the physical video cassette? A cataloger who is not certain about this may end up creating representations that reflect different entity levels within the same organization system. How does the end user deal with that inconsistency?

Once we make a decision on entity level, we can move on to specifying the metadata elements that comprise the representation.

Summary

Although the terms entity and entity level seem abstract, they provide conceptual tools for deciding what chunk of information the representations describe. An organization system must be clear on what entity level is represented, and communicate clearly to users what a single record in the system represents. Patrick Wilson discusses the five specifications for bibliographic instruments, and one of those is the unit of analysis, or entity level.

Cites & sites