Concepts of information organization

1.4. Representing information

Overview

This module introduces the concept of representation as it applies to information organization. Representation is a central concept for this course, and is a foundation for our explanation and exploration of information organization. One reason why this is such an important concept is that it allows us to explore different ways to represent the same information object for different users. One does not have to assume there is just one way to represent an object. In addition, we will see throughout the course that there are multiple levels and types of representation involved when we organize information.

Recorded information as representation

A person's thoughts, feelings, opinions, and ideas can be information to share with another. Writers record their ideas and feelings in creating an essay or book. Artists express their artistic ideas in a painting or sculpture. Musicians write scores and play instruments. All of these are instances of human expression. Are the expressions exactly what is inside the person? We have no way of knowing for sure, but we can say the expressions represent something in the person's mind.

A first phase of representation, then, is a person's expression. When the expression is recorded in a book, a painting, or a music score, we now have an instance of recorded information—an information object. In some cases, such as a musical performance or a dance, the expression may or may not be recorded. But if we record the sound of the performance on audio tape or video tape the dance performance, we now have information objects that represent those events.

Representing the information object

The role of the information organizer, the cataloger, the indexer, the metadata specialist begins at the point when an instance of recorded information exists. Information professionals follow rules and standards to create representations of information objects. The representations serve as surrogates for the objects. Users interact with these representations when they seek information through one of our systems of organization.
Let's take an example of a book. The book represents the thoughts, feelings, etc. of its creator. We can put the book on a shelf in a library and hope someone passes by and finds it! A better approach is to set up a system in which we collect representations of all the books in a collection. We provide an orderly arrangement of those representations so users can systematically search the representations to find objects that may contain the information they are interested in.

That's precisely what a library catalog is (whether in the older paper card catalog form or the current online catalogs). The catalog brings together "surrogate records," records that "stand in for" and represent the individual objects in the library's collection. Typically, each library catalog record is a representation of one object.

The challenge to the information professional is developing a representation of an information object that makes sense to the user. The challenge is to create a representation that contains the right sort of information for the user. Generally, the information professional describes and represents in the surrogate record information about the information object's container and content. We treat each of these parts of the object in separate ways, but the result is a single representation or surrogate record that describes the entire object.

**Representing the container**

The container is the physical, tangible manifestation of an information object. We begin by representing aspects of the container because this is somewhat more straightforward than representing the content.

Any physical object (e.g., a person, a house, a car) has attributes, which are characteristics, features, or properties. For example, a person has attributes of eye color, height, weight, and date of birth. The attributes provide a way to build a representation of any person, regardless of the specific values of the attributes for each person. Driver's licenses, for instance, are representations of people, and the value of the attribute height may be 6 feet for one person and 5 feet 2 for another person.

An information object has many attributes: the name of the object, a creator, a date of production, a size, a color, etc. Some of the features may be very important to represent since your users may need to know these pieces of information. These attributes become the basis for representing the container. The following example shows a representation of your textbook from the UNT library catalog:

```
Author       Taylor, Arlene G., 1941-
Title        The organization of information / Arlene G. Taylor
Edition      2nd ed.
Descript     xxvii, 417 p. : ill. ; 27 cm.
Series       Library and information science text series
Note         Includes bibliographical references (p. 385-405) and index.
Subject      Information organization
ISBN/ISSN    1563089696 (pbk. : alk. paper)
             1563089769 (hbk. : alk. paper)
LC CARD #    2003058904
```

In the example, the labels related to attributes is in bold type and the information following the label is the value in italics. In library cataloging, certain attributes of objects are part of representations and others are not. For example, in this representation there is nothing about the color of the book.

Determining the attributes of objects to highlight is critical. One important consideration in choosing what to highlight is the needs of your users. We will have more to say about this in later modules.
Representing the content

In the above example, if that were the entire representation of the book, then users would need to know the title or author in order to search the library catalog. But we know from personal experience that people look for information about certain topics, or information based on subjects. For example, I am interested in information about eradicating squash bugs. I don't know any titles or authors of books that contain that information. I need to search based on the content of the object, not its physical features.

The next step in representing information objects is to communicate to the user what the object is about, what intellectual content it contains. Many information objects, such as books, can be characterized by what they are about, what topics they cover, what themes or perspectives are presented. So, we can say that one more attribute of the object is its intellectual content or subject.

The library catalog record for our textbook also includes the following in its representation:

Subject: Information organization. Metadata.

Again, the label associated with the attribute is in bold type and the attribute value is in italics. According to the person who created the representation for our textbook, the book is about information organization and metadata. Often the intellectual content of an object is represented by subject headings, index terms, and other techniques we will discuss later. In the case of this library catalog record, the values of the subject attribute come from the Library of Congress List of Subject Headings. These two subject headings provide the user with a sense of what the books are about.

The attribute of subject itself can have more specific meanings. For example, the time period covered by the book is a type of subject. For a geography book, the geographic areas covered by the book is a type of subject. And we know that people themselves can be the subject of a book (e.g., biographies). So, subject is a multifaceted attribute.

Adding this attribute to our representation now provides the user with the ability to search for objects based on subject as well as title, author, and other attributes. Before concluding this module, we need to introduce one more concept. If you look at the listing of subjects covered by the textbook, you will see the term metadata. Along with the concept of representation, metadata is another central concept in organizing information.

Metadata and representation

The term "metadata" may be completely unfamiliar to you now, but by the end of the course, you will have a firm understanding of the concept and its uses. A later online module will explore the concept and its use in more detail.

For now, metadata can be defined simply as "data about data." As you may be beginning to realize, information organization is in some ways about providing users with information about other information! Representations in the library catalog describe objects in the library's collection. Metadata is the general concept for this. The library catalog record is a metadata record that represents information objects in the library's collection.
Metadata is structured information representing an information object. In the example of the catalog record above, we structure the representation of the attributes of the object into separate pieces of information: information about the creator, information about the title, etc. Each of these pieces of information is labeled (e.g., Author, Title, Subject). We label the information to help the user understand the representation and ultimately for computer processing of this information.

What is the relationship between attributes and metadata? An attribute is a general characteristic or features of the object. Metadata elements provide the way to label these attributes. In the above example, then, the labels of Author, Title, Subject, etc. can be called metadata elements. Metadata elements are used to label pieces of information in our representation. The library community has an entire set of metadata elements that are based on the rules for cataloging.

A metadata record is a representation of an information object. Each piece of information that our users need to know about the object is represented, and the metadata elements provides the labeling and structuring of the representation.

Summary

In this module, we discussed two central concepts related to information organization: representation and metadata. Representation allows us to think how we can best present information about an information object to our users. Through our representations, we highlight certain attributes (e.g., its intellectual content) while ignoring others that may not be pertinent to our users (e.g., its color). Another set of users may need to know about the color of an object, and the concept of representation allows us to think about what is important to represent for a particular set of users. There is always a tradeoff when we represent something. Some things are not represented. The challenge is to present the most appropriate representation for our users.

Representation gives us the pieces of information to highlight for the user and, metadata becomes the organizing principle for structuring the representation. Users and computers interact with these structured representations. Metadata records representing information objects are a critical tool for connecting users with information.