Overview

This module ties together major concepts introduced in previous basic concepts modules. We have said that representation is the central concept in information organization, and that information organization involves different types and levels of representations. We are primarily interested in two types of representations: representations of users’ information needs and representations of information objects. We have also said that it is vital to understand the behavior of information users in order to be able to create metadata representations that serve them effectively.

Here we present an overall model of information retrieval (IR) that illustrates the relationships among concepts of users, representations, and the IR system or database. This model will be used throughout the course as a point of reference in our approach to information organization.

A model provides an abstract view of a phenomenon. It simplifies components of the phenomenon to provide a picture that reveals the components and their interaction. One objective of this series of modules is to help you develop a mental model of information organization and its role in information retrieval. As you increase your understanding of the concepts during the course, your mental model will become more complete as we add detail to these ideas and they begin to fall into place.

Information retrieval model

In the overall model of information retrieval . . .

1. Representations of information objects—metadata—are entered in the information retrieval system.
2. During information retrieval, a user with an information need creates a representation of that need—a query—and enters it in the IR system.
3. The system attempts to match the query with the metadata.
4. If the match is successful, the system retrieves or displays to the user a subset of all the metadata representations.
5. Finally, the user evaluates this subset of representations, or the system’s output.

The figure below illustrates the complete process.
Important questions about IR

The model suggests several critical questions and answers to bear in mind throughout this course.

- **Which comes first, metadata or users?**

  The model shows metadata representations on the left because metadata must be put in the IR system before it can be retrieved.

  In user-centered system design, however, the process begins and ends with users.

  - First we analyze users' information needs as represented in their questions. From analysis of both users' questions and actual information objects, we develop a way to structure metadata representations of information objects to put in an IR system.

  - After the IR system is implemented and users retrieve metadata from it, we determine how well the metadata serves the users. That is, we evaluate the effectiveness of the metadata in order to be able to improve the system.

  Therefore, the metadata representations that we create are both based on and respond to users' information needs.

You will get hands-on experience with user-centered design in your information organization project.
• **How does the system match the representations?**

In the figure above, the IR system is shown as a black box. This is a visual pun in that "black box" is a popular expression for a device that works in some mysterious or unknown way. An IR system works through matching words (data) in the queries with words in the metadata. Later we will explain the principles of just how the matching occurs. The most important point to remember now is that computers (unlike people) cannot deal with ambiguity, so the match must be **exact**.

• **How does the user evaluate the retrieved subset of metadata representations?**

The user's evaluation process, called relevance evaluation, is by definition individual and subjective, depending on the user's particular situation. Our responsibility is to create metadata representations that allow the user to make adequate relevance judgments. Can the user make a judgment about the object based on its representation? Are the representations appropriate, sufficient, and accurate? If users cannot make sense of the representations, they may not be successful in finding the information objects that hold answers to their questions.

• **When is the user's information need resolved?**

We have said that an IR system allows the user to search for and discover information. In fact, most IR systems allow the user to search and retrieve only metadata representations of information objects. It is important to recognize that the IR system does not resolve the user's information need. Instead, the metadata representation in the system points to an object that may hold the information needed by the user. The user must obtain the actual information object (another level of representation) and then decide whether that object meets his/her information need.

It is also important to remember that in this course we focus on IR systems as closed databases and that we assume a need for metadata. In the broadest view, not all IR employs metadata (e.g., on the World Wide Web), a situation that we discuss only briefly. Other courses cover information access on the Web in depth.

**Summary**

In this module, we begin to connect the concepts introduced so far. We indicate two major types of representations in an IR model that illustrates the central role of representations in organizing information for retrieval. We point out that the IR system does not provide the information that directly answers the user's information need, but rather provides metadata representations that point toward potentially relevant information objects. This important function enables users to get that much closer to finally being connected to the information they need. Finally, in explaining the IR process, we underscore the importance of understanding the users who are to be served by the IR system.