Library of Congress Subject Headings
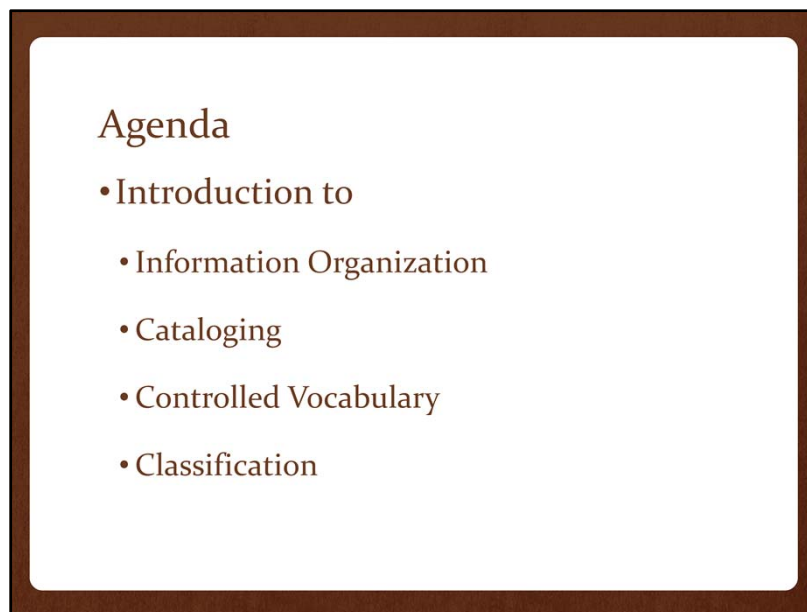
Module 1.1:
Why Do We Organize Information?

Policy and Standards Division
Library of Congress
June 2016

> **Agenda**
> - Introduction to
>   - Information Organization
>   - Cataloging
>   - Controlled Vocabulary
>   - Classification

In this first module of the training series, we begin by reviewing some concepts related to *information organization* before jumping directly into subject cataloging with *Library of Congress Subject Headings* and Library of Congress Classification.

We discuss the role of information organization, cataloging, controlled vocabularies, and classification.

In later modules, we will discuss more specific aspects of subject representation using Library of Congress standards.

We will begin with the overview of Information Organization.

Information Organization is a process of establishing systems so that documents and information can be retrieved by users whenever needed.

It is a process of describing information resources (for example, books, web pages, databases, and so on), and making them findable by identifying and recording names, titles, and subjects associated with each resource.

These descriptions may take on a variety of forms. For example, they may be formatted as comprehensive records that describe the resource. Some descriptions may be embedded in the documents themselves, especially for online resources. And others may be created as simple, discrete metadata statements stored on the Web.

These descriptions act as pointers to the actual information resource (by providing location information, such as a call numbers, or direct links to documents).

Information Organization is part of the larger library and information science landscape. It is one of the key services provided in libraries, archives, museums, digital environments, and so forth. It is sometimes referred to as *bibliographic Control*. It includes more specific activities, such as cataloging, indexing, archival description, and all other forms of metadata creation.

In this training, we will focus on library cataloging.
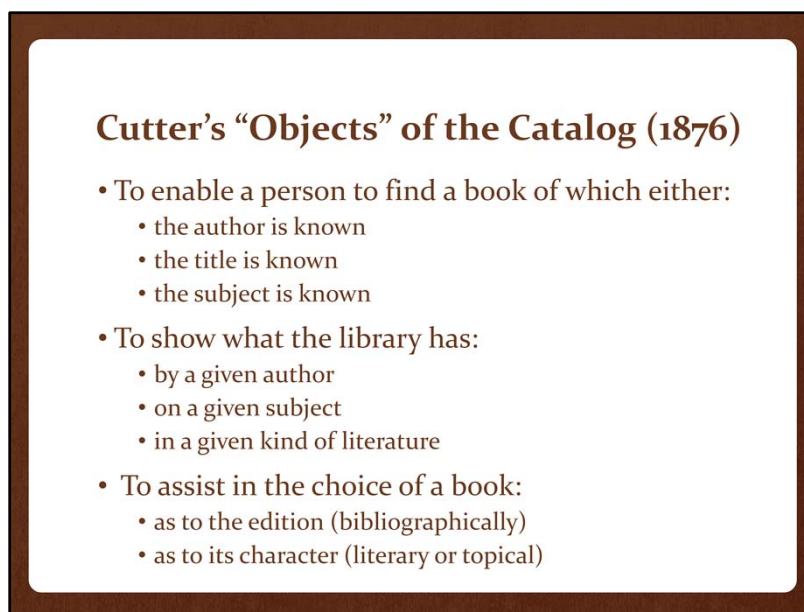
Why do we organize information?

- Life, documents, collections, and institutions are messy ...
  - and, our personal approaches do not work at an institutional level.

- Information Organization is about ...
  - retrieval
  - collocation
  - sharing
  - understanding

Without a system of organization, life, collections of resources, and institutions can get very messy. We organize information to store it, retrieve it, relate it to other things, and to understand it.

The larger a collection grows, the more difficult it can be to find things quickly and efficiently. And the personal approaches we take with our own information and our possessions (for example, organizing by color or by size) they do not scale well to handle larger collections. One cannot use ad-hoc, informal strategies in information institutions to support the retrieval need of *others*.

In order to provide access to others, it's important to carefully plan how information will be organized. And, we do this by using standards.

A shared vision of *information organization* allows us to make retrieval more predictable, to bring together like resources on the shelves and in our retrieval tools, to share information with other institutions, and to understand the nature of our resources, collections, and the other entities found in our catalogs (such as names, titles and subjects).

Cutter's "Objects" of the Catalog (1876)

- To enable a person to find a book of which either:
    - the author is known
    - the title is known
    - the subject is known
- To show what the library has:
    - by a given author
    - on a given subject
    - in a given kind of literature
- To assist in the choice of a book:
    - as to the edition (bibliographically)
    - as to its character (literary or topical)

The shared vision of information organization has been guided by principles. One of the most influential is Charles Cutter's "Objects of the Catalogue," first published in 1876 in his *Rules for a Printed Dictionary Catalogue*. He discussed what we expect our retrieval tools to do. If we look closely, we can see that subject and genre/form access were considered by Cutter to be important functions of the catalog.

The first object (or Objective) is to be able to find a resource if the name of the creator, or the title, or the subject of the resource is known. This finding function insists that the catalog provide access to individual resources in a collection.

The second object is the collocation or gathering function. It states that users should be able to retrieve not only a single resource, but to discover all resources related to a creator, a subject, or in a particular genre or form of material.

The third objective reflects a selection function. It means that users should be able to choose a resource that best fits their needs, based on either its bibliographic characteristics (such as language or format), or as to its subject-related or genre/form characteristics.

Cutter's definition of the catalog has stood the test of time, but it has been updated and expanded over the years.

> ## IFLA's *Statement of International Cataloguing Principles* (2009)
>
> - From International Federation of Library Associations and Institutions (IFLA)
> - Identifies user tasks based on IFLA's *Functional Requirements for Bibliographic Records* (FRBR; 1998)
>   - to **find**
>   - to **identify**
>   - to **select**
>   - to **acquire** or **obtain**
>   - to **navigate**
> - Undergoing revision 2015/2016

Most recently, after years of collaboration and discussion among cataloging experts from around the world, new international principles were published in 2009 by IFLA (The International Federation of Library Associations and Institutions). In these principles, the functions of the catalog are to enable a user

- to **find** bibliographic resources in a collection . . .

- to **identify** a bibliographic resource or an agent, which is an entity that is responsible for the resource . . .

- to **select** a resource that is appropriate to the users' needs . . .

- to **acquire** or **obtain** access to the resource. . .

- and to **navigate** within a catalog and beyond.

These user tasks are based on IFLA's *Functional Requirements for Bibliographic Records*, also known as "Ferber" or FRBR. FRBR is an abstract conceptual model of the bibliographic universe.

In order to meet IFLA's user tasks, we, as catalogers, must create adequate and well-formed metadata!

What is *Metadata*?

- Data about data

- "Structured information that describes the attributes of resources for the purposes of identification, discovery, selection, use, access, and management." – Joudrey, Taylor, and Miller (2015). *Introduction to Cataloging and Classification*, 11th ed. Santa Barbara, CA: Libraries Unlimited.

What is metadata?

The best known definition is the easily remembered "Data about Data."

Other definitions, however, are much more descriptive.

In the 11th ed. of *Introduction to Cataloging and Classification*, Joudrey, Taylor, and Miller state that metadata is: "Structured information that describes the attributes of resources for the purposes of identification, discovery, selection, use, access, and management."

In short, metadata is information about a resource. This broad definition includes elements such as titles; edition statements; the names of creators, contributors, and others; subjects; dimensions; location information; contents; and so on.

Metadata allows users to find, identify, select, and obtain the resources in our collections.

What is *Cataloging*?

- "The process of creating metadata for resources by describing a resource, choosing name and title access points, conducting subject analysis, assigning subject headings and classification numbers, and maintaining the system through which the cataloging data is made available." – Joudrey, Taylor, and Miller (2015). *Introduction to Cataloging and Classification*, 11th ed. Santa Barbara, CA: Libraries Unlimited.

In libraries, metadata creation is often called cataloging.

Cataloging is a subset of the larger field called *information organization.*

It can be defined as, "The process of creating metadata for resources by describing a resource, choosing name and title access points, conducting subject analysis, assigning subject headings and classification numbers, and maintaining the system through which the cataloging data is made available."

In other words, it is the process of creating metadata about library resources, which is included in the catalog.

A *catalog* is an organized compilation of bibliographic metadata that represents the holdings of a particular institution or a library network and/or resources accessible in a particular location.

Cataloging, traditionally, has been seen as comprising two specific activities: *Descriptive Cataloging* and *Subject Cataloging*.

Descriptive Cataloging

- Identification of the resource

- Description of resource

- Access points for names and titles

- Authority work for names and titles

Descriptive cataloging is that phase of the cataloging process that is concerned with the identification and description of an information resource, answering questions such as:

- What is it?
- What are its distinguishing and significant characteristics?

It is also concerned with encoding this information for machine processing, for example in:

- A MARC record,
- An XML encoded description, or
- A metadata statement compatible with linked data principles.

It is also concerned with the selection of names and titles useful for providing access to the resources and the establishment of authorized access points for names and titles, answering questions such as:

- What is it called?
- Who is responsible for its creation?
- Who else contributed to it?
- By what form of name or title are they best known?

Descriptive cataloging describes the makeup of an information resource and identifies those entities responsible for its intellectual and/or artistic contents *without* reference to its classification by subject or to the assignment of subject headings, both of which are the province of subject cataloging.

Subject cataloging is the phase of the cataloging process which is concerned with determining and describing the intellectual or artistic content and the genre/form characteristics of a resource, and translating that understanding into subject headings and classification notations.

After the resource's *aboutness* has been determined, as many subject headings as are appropriate are chosen from a standard list. There are many such lists, including: *Library of Congress Subject Headings* (known as LCSH), the *LC Genre/Form Terms for Library and Archival Materials*, *Medical Subject Headings* (which is often referred to as MeSH), the *Art and Architecture Thesaurus*, and so on.

In addition, a classification notation is chosen from whatever classification scheme is used by the library. In the United States, the most likely candidates are the Library of Congress Classification (often referred to as LCC) or the Dewey Decimal Classification (known as DDC).

Traditionally in the U.S., the classification serves as a means for bringing a resource into close proximity with other resources on the same or related subjects. In the case of tangible resources, the classification is the first element of the call number, which is a device used to identify and locate a particular resource on the shelves.

In later modules, you will learn much more about these topics.

For consistency and improved retrieval, libraries and other information institutions attempt to suppress the anarchy of natural language when it comes to describing the aboutness of resources. Subject cataloging is more consistent when the vocabulary that is used is controlled.

The main objective of vocabulary control is to promote the consistent representation and comprehensive searching of subject matter.

This is achieved through the control of synonymous and nearly synonymous terms, by distinguishing among homographs and homonyms, and by linking together terms whose meanings are related in some fashion (identifying broader, narrower, and related terms).

But the use of controlled subject languages is only part of subject cataloging; the other part involves classification.

What is *Classification*?

- The use of a system of notations or symbols to categorize the contents of resources
  - Used to provide logical shelf arrangement (i.e., call numbers), but also can be useful in searching the catalog
  - Organized by disciplines, they begin with broad general topics which give way to more specific subtopics

According to Joudrey, Taylor & Miller, classification is "the placing of subjects into categories."

It is the process of determining where a resource fits within the classification's structure, and then assigning the notation that most closely approximates the aboutness of the resource.

Classification is more than finding the right notation or category; it is about relationships. It provides a logical arrangement of topics and subtopics from the general to the specific that can be translated into a linear arrangement for materials in a library. Classification traditionally provides formal, orderly access to the shelves, but it is also a mechanism by which to collocate materials in the catalog. It's what makes browsing possible.

Both of these tools, controlled vocabulary and classification, provide different methods of subject access to the materials in our collections.

Classification and controlled vocabularies are discussed further in other modules.