### Knowledge Organization

- **Is there a basic human need to put things into some sort of order?**
  - Much of natural language concerns categories of things rather than individual things
  - Why do we organize things and information?
    - Why do spoons go in THAT drawer in the kitchen and not in a can in the garage?
    - Why do your favorite books go on one shelf and not-so-favorite on another?

### What is a Library?

- **Is a library simply a building filled with books?**
  - Let’s examine this question from an organization of knowledge perspective, i.e., how are libraries organized so that people can find the information that they need?
  - Technique #1: Use of known call numbers: People understand that books are stored in libraries using call numbers
  - Technique #2: “Serendipitous” discovery: People understand that books are stored in libraries in a way that groups books together by subject
  - Technique #3: Use of online catalog (OPAC): People understand that a catalog is available for search
  - Our examination will require that you think analytically about your past library use (those uses that have been both successful and unsuccessful).

### What is a Library – Analyzed

- **Recall an experience when you walked into a library and found the paragraphs that you needed in various non-fiction books:**
  - Found front door.
  - Located OPAC and determined call numbers.
  - Followed signage to stacks.
  - Located books.
  - Used table of contents or book index to locate needed paragraphs in each book.

### The “Great Reduction Phenomenon”

- Experienced users can walk into any professionally-managed library and locate or relocate specific information. Why?:
  - All libraries share a basic architectural identity:
    - Physical
    - Conceptual
  - All non-fiction books share a basic architectural identity:
    - Tables of contents
    - Book indexes
The National Network of Librarians

– All professional librarians share their work:
  • Why?: “Library economy”
  • How?: Standardized practices, such as centralized cataloging by the Library of Congress
– Librarians from conceptually similar libraries share their work:
  • Why?: “Library economy”
  • How?: Standardized practices

The Great Reduction Phenomenon

• It takes a profession.
• A profession with standardized practices.
• And preferably, standardized information packages:
  – Books and articles: highly standardized structural characteristics across publishers
  – Websites (also call networked resources): ????

Knowledge Organization

– Examination of organization of knowledge techniques:
  • Across different types of information packages, including print and networked resources
  • Across different practice environments, including public, research, and special libraries
– Our point of view: information storage and retrieval:
  • What are we storing: structured surrogate records representing information packages of various types
  • How are we storing: cataloging and metadata content rules

Libraries as our Focus

• The predominant types of information packages in library collections are highly standardized.
• The organizing practices performed by library catalogers are the result of over 125 years of professional activity and standards development:
  • Left side: what data elements are required
  • Right side: rules for surrogate record data entry

Example: Online Public Access Catalog

• Catalog of what?: What is held by a library.
• What is stored in the catalog?: Bibilographic records representing books, sound recordings, films, computer files, maps, realia ....
• How are bibliographic records stored? The topic of our course.

General Steps for Organizing

• Obtain information package.
• Describe information package.
• Perform subject analysis on information package:
  • Verbal analysis (indexing)
  • Classification
Why Organize Knowledge?

- The main reason
  - So that you can find things more effectively
- I.e., effective retrieval is predicated on some sort of organization applied to information resources
- Historically there have been many institutions and tools devoted to knowledge organization
  - Libraries
  - Museums
  - Archives
  - Indexes and catalogs, dictionaries, phone books, etc.

Why Organize Knowledge?

- A question of scale
  - Using your own ad hoc set of categories and methods to organize your own collection of books or CDs seems to work fine...
  - What if your collection grew to
    - 10 Times the size? How would you organize it?
    - 100 Times?
    - 1000 Times?
    - 100000 times?
    - What if it wasn’t physical objects, but electronic?

What is Knowledge Organization?

- Identifying the existence of all types of information-bearing entities as they are made available
- Identifying the works contained within those information-bearing entities or as parts of them
- Systematically pulling together these information-bearing entities into collections in libraries, archives, museums, Internet communications files and other such depositories

What is Knowledge Organization?

- Producing lists of these information-bearing entities prepared according to standard rules for citation
- Providing name, title, subject and other useful access to these information-bearing entities
- Providing the means of locating each information-bearing entity or a copy of it

Goals for this Course

- Equip you with the necessary knowledge of how retrieval tools for print and networked resource collections are organized so that you can be an effective information services or reference librarian.
- Provide basis for further specialized study in knowledge organization areas
- This is NOT a cataloging class per se, but a course on different retrieval tools to determine how surrogate records were created in order to see similarities and differences across these tools, i.e., to help you become an information professional.

What is Bibliographic Control?

- THE ART OR SKILL OF ORGANIZING KNOWLEDGE FOR RETRIEVAL
  - KNOWLEDGE: WHAT IS KNOWN
  - INFORMATION: KNOWLEDGE PERCEIVED
- THE PROCESS OF CREATING, STORING, MANIPULATING, AND RETRIEVING BIBLIOGRAPHIC DATA
  - BIBLIOGRAPHIC DATA: INHERENT
  - BIBLIOGRAPHIC CHARACTERISTICS
**MODES OF ORGANIZATION OF KNOWLEDGE**

- Indexers/Abstractors: primarily journal articles, essays in anthologies, specific formats of materials (like dissertations, newspaper articles) -> description, summarization
- Catalogers: historically books, journals, etc., but also music and other materials to be found in a library collection -> description, summarization (highly standardized)
- Archivists: primary records and documents, levels of description and summarization
- Historical bibliographers: rare books -> more exhaustive textual description and annotation
- Art curators: art materials -> varying, idiosyncratic methods of description
- Systems management, computer professionals: electronic information, either databases or processes -> information processing, structural algorithmic manipulation

**TYPES OF BIBLIOGRAPHIC CONTROL (RETRIEVAL TOOLS)**

- **INDEXING**: controls periodical literature of a discipline
- **CATALOGING**: controls the collection of a library or group of libraries

**CHARLES AMNIE CUTTER (1876) PURPOSES OF A CATALOG**

1. TO ENABLE A PERSON TO FIND A DOCUMENT OF WHICH THE AUTHOR, OR THE TITLE, OR THE SUBJECT IS KNOWN
2. TO SHOW WHAT THE LIBRARY HAS BY A GIVEN AUTHOR ON A GIVEN SUBJECT IN A GIVEN KIND OF LITERATURE
3. TO ASSIST IN THE CHOICE OF A DOCUMENT BIBLIOGRAPHICALLY (E.G. EDITION, FORMAT, ETC.) AS TO ITS CHARACTER (I.E. LITERARY OR TOPICAL)

Professional Librarians have Created:

- A worldwide network of libraries:
  - In geographic neighborhoods
  - In conceptual neighborhoods
- Libraries with sets of retrieval tools available
  - OPACs
  - Bibliographic databases
  - Full text databases
  - Bibliographies
  - Web resource collections and search engines

**What is Stored in Retrieval Tools?**

- Surrogate records.
- In professional tools, however, they are **structured** surrogate records.
- But, what do we mean by structured?

**Structures**

- The architectural elements of a house
- The anatomy of the human body
- Common structural elements of surrogate records – the cards of the card catalog:
  - Card catalog structure includes following elements:
    - Author
    - Title
    - Subject
Surrogate Record Structure

- The left side (element set) provides access points:
  - **TI** = All about Lovebirds
  - **AU** = Soderberg, Pearcy Measday
  - **SU** = Lovebirds
- The right side (content) provides searchable fields:
  - **TI** = All about Lovebirds
  - **AU** = Soderberg, Pearcy Measday
  - **SU** = Lovebirds

Some Definitions and their Contexts

- Bibliographic data – at least three contexts
  - The content of professionally managed library surrogate record collections (e.g., library catalogs)
  - The content of professionally managed information products (e.g., article databases)
  - The content of professionally managed archival surrogate record collections (e.g., archival finding aids)
- Metadata – at least two contexts:
  - Specific definition: Non-traditional bibliographic data
  - General definition: Any information that represents other information (including cataloging, indexing, etc.)

Professionally Managed?

- Professionally managed metadata consists of collections of structured surrogate records.
- In library cataloging, we have standards for what element sets to employ (left side) and rules for how to consistently enter content into each record (right side).

Professional Use of Retrieval Tools

- Know what type of “information package” is represented by the surrogate records of a retrieval tool – if you are seeking:
  - Cited articles – then consult a bibliography
  - Library items – then search a catalog
  - Published articles – then search a bibliographic database
  - Archived items – then use an archival finding aid
  - “Published” web pages – then a search webpage database
- Understand underlying structure of the surrogate records of retrieval tools:
  - How they are similar
  - How they differ

Professional Use of Library Catalog

- Know that a library catalog contains surrogate records for a variety of types of “information packages:”
  - Books
  - Maps
  - Computer software
  - Electronic resource
  - Realia
- Understand that the structure of the surrogate records for each different types of information packages will vary

What about use by non-professionals?

- Lazy users?
- Complex retrieval tools!
- Professional’s role is to mediate or to teach about the effective use of retrieval tools by end users.
Jargon Check

- [Bibliographic] control – the systematic representation of information packages in surrogate form for purposes of their effective and efficient retrieval.
- BIBLIOGRAPHIC CONTROL: THE PROCESS OF CREATING, STORING, MANIPULATING, AND RETRIEVING BIBLIOGRAPHIC DATA
- Surrogate record – think about the cards of the card catalog.
- Collocation – the bringing together, or retrieval, of surrogate records according to a criterion. Think about the various drawers of the card catalog: author, title, subject.

Review of Basic Retrieval Tools

- How are surrogates structured? (information storage)
- How are structured surrogates displayed for efficient (and transparent!) perusing? (information retrieval)

General Purpose of Retrieval Tools

- To provide enough information for end users to determine whether or not to pursue the acquisition of actual information package.
- This is a question of (library) economy – save the time of the user – whether the information package is a book, an archival artifact, or a webpage.

Bibliographies

- List of surrogates
  - Appendix lists (e.g., scholars who cite papers)
  - Compiled lists (e.g., free-standing bibliographies)
  - Inventory lists (e.g., bookstores)
- Structured surrogates that are collocated according to a criterion.
- Facilitate transparent use – browsing surrogate list.
- Provide key information for uniquely identifying targeted information packages.

Bibliographies – Appended Lists

- Citation is criterion – most common type of bibliography:
  - Cited scholarly articles
  - Cited books and other types of information packages
  - Style guides serve as standardizing display mechanism over time:
    - Across authors
    - Across editors
    - Across journals
    - Across disciplines

Bibliographies – Compiled Lists

- Independent works
  - Found in catalogs (e.g., A Bibliography of Austin Dobson)
  - Often the result of scholarly labor (i.e., used for tenure), including annotated bibliographies
- Collocated according to various criteria:
  - Subject bibliographies
  - Author bibliographies
  - Bibliographies of a time period
  - Language, location, publisher, form bibliographies
Indexes to the Literature

Collection of surrogate records representing the analyzed contents of information packages:
- Journals
- Conference proceedings
- Chapters
- Web sites
- Have additional left side elements (i.e., access points):
  - Familiar access points include: author, title, subject
  - Other access points, to be discussed, that are possible because of the computerization of indexes

Index Interfaces

- Print indexes resemble print bibliographies
- Online indexes (also called bibliographic databases)
  - HW Wilson
  - Library Literature
  - General Science Index
  - EBSCO
  - PsychInfo
  - MLA International Bibliography
  - LitFinder

Finding aids

Collection of surrogate records that present long descriptions of archival collections:
- Often at the box level
- Occasionally at the item level

Library catalogs

Collection of surrogate records representing the collection of a single library (cf “union catalog” and WorldCat).
- Have multiple access points (i.e., how people would ask for an information package):
  - Traditional (i.e., from the card catalog era): author, title, subject
  - Other access points, to be discussed, that are possible because of the computerization of the catalog

TYPES OF CATALOGS: Book Catalog

- Replaced by cards
- Portable and can be located and consulted in any part of the library
- Multiple copies easily made
- Entries are grouped so the eye can scan a page quickly in contrast to turning one card at a time
- Occupies less space than card catalog
- Clerical errors (filing) can be eliminated if arrangement done by computer

TYPES OF CATALOGS: Card Catalog

- Popularized in USA by Library of Congress in 1901
- New entries can be added at any time; therefore the catalog can be updated constantly
- Entries (a representation of a bibliographic record at any point in an index or catalog) are on movable units (cards) so new entries can be interfiled among the old. This means that subsequent searching is done in only one place (the same tray), not in several supplements
- Cards can be removed at any time, making it possible to withdraw entries immediately upon discarding materials. Cards can be changed or replaced. The card catalog can keep pace with changes in forms and headings
- The alphabet is distributed over many trays enabling more users to use it at one time
- Print is larger and easier to read
- More information appears on the card than in book catalogs
### TYPES OF CATALOGS: COM Catalog

- Constructed like a card catalog; produced like a book catalog
- Same advantages as book catalogs, if film readers are available

### TYPES OF CATALOGS: Online Public Access Catalog (OPAC)

- No delays from filing and card production
- Instant updating possible
- Global updating possible
- All entries in a single sequence like card catalogs; no supplements
- Terminals available everywhere (dedicated or dial)
- Much more information can be given without the physical constraints of other types of catalogs. Much more flexibility in search strategies

### TYPES OF CATALOGS: Web Based OPACs

- Hypertext links
- Familiar user interface

### TYPES OF CATALOGS: Alphabetical or Classified

- **Advantages of alphabetical catalogs**
  - Alphabet is commonly understood
  - Reader is likely to find information under direct terms (those consulted first by choice)
  - Verbal headings are more easily understood than classification symbol
- **Disadvantages of alphabetical catalogs**
  - Rapid obsolescence of terms
  - Multiple meaning of terms
  - Growth of catalog increases its complexity
  - Interfiling different types of entries may confuse the user (fail to distinguish between title and subject entries)

- **Advantages of classified catalogs**
  - Notation is independent of language. Obsolescence is not a problem. International standardization is easier
  - Closely related classes brought together in sequence; better for scanning or for moving from general to specific
  - Permits easy compilation of subject bibliographies
- **Disadvantages of classified catalogs**
  - Order of arrangement is uncommon; unlike alphabet
  - Subject index is a necessity and must always be used
  - When a system of knowledge is altered, the entire scheme must be replaced
  - Notation system may create impression of intricacy (psychological barrier to use)

### TYPES OF CATALOGS: Dictionary vs. Divided

- Dictionary catalogs have one sequence which contains author, title and subject entries interfilered.
- Divided catalogs attempted to break up large files in dictionary catalogs. Author-title/subject or Author/Title-subject or all three separately
- **Advantages of divided catalog**
  - Save filing time and money
  - Filing easier for user to comprehend
- **Disadvantages of divided catalog**
  - May increase user confusion. Adds extra steps
  - Confusion in dividing the entries (author-title subject entry for the criticism of a work? Where?). Requires duplicate references
Parts of the Catalog

1. PUBLIC CATALOG
   ALPHABETICAL OR CLASSIFIED
   DICTIONARY OR DIVIDED
2. SHELFLIST
   ENTIRE COLLECTION IN SHELF ORDER
   HOLDINGS LOCATIONS
3. OFFICIAL CATALOG
   ONE OF EACH CARD
4. AUTHORITY FILES
   RECORD OF ALL ACCESS POINTS

OBJECTIVES/QUALITIES/REQUIREMENTS OF A CATALOG

• UP-TO-DATENESS
• FLEXIBILITY
• LOGICAL STRUCTURE
• EASE OF USE
• ECONOMICALLY MAINTAINABLE AND OVERALL COST EFFECTIVENESS
• COMPACTNESS/PORTABILITY/REMOTE ACCESS/COPIABLE

INFORMATION RETRIEVAL PROCESS

THREE STAGES:
1. SUSCEPTIBLE PERSON WITH INQUIRY
2. EXPRESSIONS OF INQUIRY IN THE LANGUAGE OF THE RETRIEVAL SYSTEM
3. SET OF RETRIEVED SIGNALS

FUNDAMENTAL COMPONENTS OF THE INFORMATION RETRIEVAL PROCESS
1. FORMULATION OF THE INQUIRY
2. RETRIEVAL OF SIGNALS
3. UTILIZATION OF WHAT HAS BEEN RETRIEVED

Why do we do it?

• THE PURPOSE OF BIBLIOGRAPHIC CONTROL IS TO AID IN HUMAN PROBLEM SOLVING