

# Systems integration in libraries

LIBR 509  
Week 10 Mar 21 2022

## 1 Recorded Lecture

### 1.1 Review of Systems

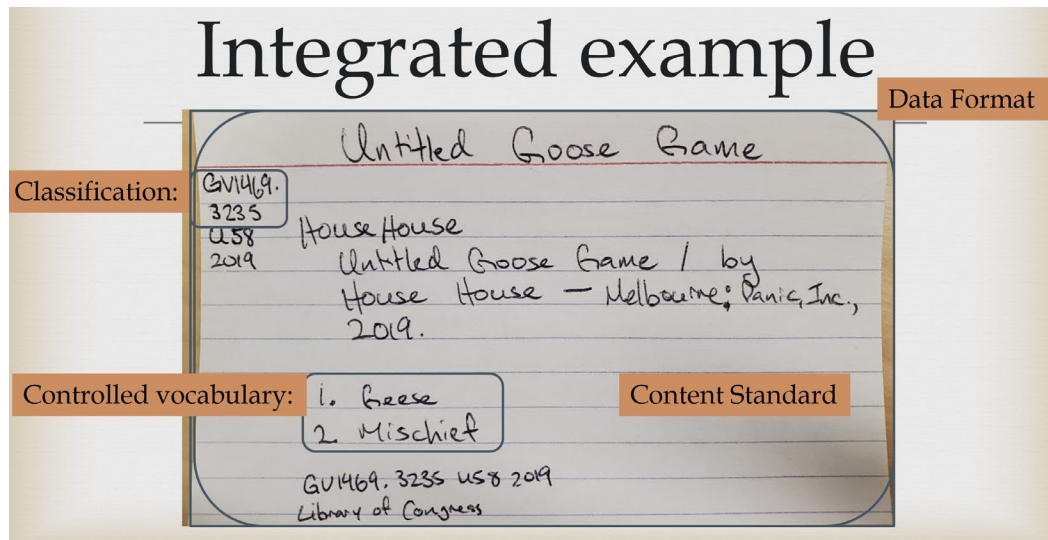
So far in this course we have been talking about a lot of systems in isolation:

- **Classification schemes** (hierarchical and faceted; ways in which the shelf order of print materials are decided; how large groups in the library are delineated)
- **Controlled vocabularies** (mostly in the context of subject headings getting at the “aboutness” of a work; also any time we have a fixed list of terms that is appropriate to describe a single element of a work—e.g., file type)
- **Content standards:** what types of things are important to describe in an item? How do we have consistent, high-quality descriptions? How do we handle discrepancies?
- **Data formats** (where will all this information go, where will it be stored, what does it mean for the user’s experience of navigating thousands of records at once)

There are also a number of things that we haven’t discussed (but still relevant about what goes into a catalogue record):

- **Unique Identifier Registries (URIs)** (how do we create a unique code for every work, every person, every record)
- **Library consortiums and copy cataloguing databases:** for most people who work with these records, the majority of the work they do isn’t original cataloguing, but rather they usually start from a partial or nearly-complete record created by another institution (particularly for libraries; less the case for archives or museums, that are likely to work with unique items)
- **Cataloguing software** (affects how those records are interpreted and made accessible; often designed with data format in mind; mediates our experience of those raw data formats)

## 1.2 An integrated example: Untitled Goose Game



- **Classification:** LoC classification number (“G” and “V” are saying something about broad category of the work; gets more specific with U58 2019)
  - Classification always in the same place on the card for easy reference
  - Library of Congress Classification
  - Consulted volume 21 LCC scheme to find entries for games and amusements / indoor games and amusements / video games / Nintendo games / individual games
- **Controlled vocabulary** being used for aboutness (subject headings) listed in order of their relevance to the item
  - Consulted controlled vocabulary for terms at right level of specificity; match against the game’s subject based on marketing, reviews, summaries
- **Content standard** being used to determine what to say about this item (how should we represent creator? Are we faithfully transcribing from item or doing some interpretive work?)
  - Here following Anglo American Content Rules 2<sup>nd</sup> edition (AACR2)
  - Transcribe information from the item using rules that assume a print resource (book or serial) and adapt as necessary using edge-cases given in the enormous volume of AACR2
- **Data format:** the card itself
  - Index card
  - Get a fresh index card, load into a typewriter with custom settings for the margins and sections optimized for constraints of index cards and AACR2 conventions; thread it into the catalogue drawer using a hole in the card
- OR: see if there is a centralized cataloguing record at Library of Congress and make a copy, add local information as needed

### 1.3 What does this all mean?

How does all of this impact the user's finding, identifying, selecting, and obtaining the item they need?

Compared to OPACs, how does this integrated system help you

- find an item ("known item search", when you know its title, author, main topic?)
- Distinguish between similar items based on their informational content or their format / materiality?
- Get you access to the item itself?

## 2 In-Person Lecture

### 2.1 Clarification Questions

#### 2.1.1 ePortfolio questions and answers

We are encouraged to put all the work we did in 509 on the ePortfolio.

Dr. Bullard recommends that we have a reflection narrative of "here's what I did in the course; which parts were easier, which parts took more work"

- Recommended that this reflection page be the link that we provide as our portfolio submission

**Make sure** that you somehow whitelist Dr. Bullard so she can see your portfolio (by default everything will be private). This can be done in various ways (e.g., make the whole thing public viewable, allow access via a password, add Dr. Bullard's email address as authorized viewable)

- Note: "Learning Significance" is the section of the portfolio post that is the easiest to hide (therefore it's okay to be honest / flawed in this section)—e.g., "I didn't understand this assignment at all"
- Important questions to answer: **what did you learn? What is the relationship between what you put here, and what you are capable of?**
- There are many ways to organize (e.g., use tags, put things under "courses"), as long as they are all viewable these are fair game.

#### 2.1.2 Exercise

The class went through the exercise of placing items in a "universe of things" (classifications, controlled vocabularies, content standards, data formats, "other stuff") on a timeline 1875-2022

Some generalizations:

- Early things were mostly classification systems (organizing systems were about supporting serendipitous discovery—people would go to the stacks looking for something and find related things, important to make sure shelf order would support search and discovery)

- Then, more work was in providing more access points for works, understanding that we are in a digital environment, assigning appropriate terms to things (**controlled vocabularies**) to maximize precision and recall
- Then, books aren't the only things we should be describing carefully (**content standards**)
- Since then, trying new things in **data formats**: not just information we are sharing it, but how are we sharing it? Can computers make leaps in logic for us? Can we make the system distributed (blockchain)

## 3 Readings

### 3.1 Required: DoO Chapter 11

When designing an information system, lots of decisions flow from initial decisions about **domain**, **scope** (breadth of resource), and **scale** (number of resources)

Larger collections need more people to organize and maintain them, creating **communication and coordination problems** that grow much faster than the collection

**Standardization** is the best way to prevent problems of scope and scale.

Organizing systems in the same domain and with nominally the same scope can differ substantially in the resources they contain and the interactions they support if they have **different categories of users**.

Designers who recognize that their systems have real consequences for people should commit to measures of **transparency** and an ongoing process of **negotiation** that enables those affected to voice concerns related to any detrimental effects the technology might have on them and their communities.

For most organizing systems other than personal ones, the set of interactions that are implemented in an organizing system is strongly determined by **economic factors**.

An essential requirement in every organizing system is ensuring that the **supported interactions** can be discovered and invoked by their intended users.

**Automated and computerized processes** can create the resource descriptions in an organizing system and their use is primarily driven by scale.

Organizing principles depend on resource descriptions, so requirements for the former are always **intertwined** with those for the latter.

Overly customized and inflexible resource descriptions or arrangements cannot easily accommodate the future growth of the collection.

#### 3.1.1 Designing and Implementing an Organizing System

Being explicit about an organizing system's requirements, scope, and scale helps with implementation:

- Make sure the system is ready to incorporate the full scope and scale of the resources

- Make sure you separate design principles from their implementation

**Architectural thinking** leads to more **modularity and abstraction** in design, making it easier to change an implementation to satisfy new requirements or to take advantage of new technologies or procedures.

- Purpose of architectural thinking: separate design issues from implementation issues, to make the system robust and flexible
- E.g., for physical books, arrange them with extra space on the shelves so there is room to add new objects to the collection
- For **digital resources**, inexpensive storage and high bandwidth have largely **eliminated capacity as a constraint** for organizing systems, with an exception for **big data** (a collection of data that is too big to be managed by typical database software and hardware architectures).

Separation of **access** vs. **ownership** (can be a cultural shift for some institutions)

Technology change is inevitable; disruptions from changes can be mitigated by standards (in data formats, description vocabularies and schemas, classification systems)

- Abstracting away from implementation is good practice (makes the system more comprehensible to external users)

### 3.1.2 Operating and Maintaining an Organizing System

How likely are resources to change over time? Expand or diminish? Does the system need to be ready to accommodate more? Allow some to leave?

The organizing system will also require maintenance over time:

- The most predictable maintenance activities for an organizing system with an expected long lifetime are **incremental changes in description vocabularies and classification schemes**.
  - Basic principles of the organizing system are not affected, but more properties are needed to maintain distinctions between types of resources
  - E.g., bibliographic classification schemes need room for new subcategories for new fields of knowledge
  - E.g., revisions to ACM keyword classification system for published articles
- Another very predictable type of activity over time with organizing systems is a **technology upgrade** that improves its quality or capabilities without affecting the organizing principles.
- The most **challenging** kinds of maintenance activities for organizing systems involve **changes to the principles for arranging** resources along with changes in the implementing technology.
  - E.g., introducing semantic web and linked data concepts in bibliographic organizing systems
  - It might be possible to roll out changes incrementally, or might need to be “all or none”

What resources are being organized? Why are the resources being organized? Who does the organizing? When are the resources organized? Where are the resources organized? How much are the resources organized?

### **3.2 Pick a Case Study from Do0 Chapter 12**

### **3.3 Suggested: McCloud, S. (1994). Chapter 2: The Vocabulary of Comics. In Understanding Comics: The Invisible Art. New York: Kitchen Sink Press.**

(this is an introduction from a very different field of how to think about representation and abstraction; if you're wanting to return to the big ideas of this course and understand them in a new way, this is a great starting place)