THE DISCIPLINE OF ORGANIZING:
The Intellectual Intersection of
the Information Schools

Robert J. Glushko
glushko@berkeley.edu

3 October 2012

Today’s Talk

- Project Motivation
- The “Organizing System”
- Design Patterns for Organizing Systems
- The Book, eBooks, Customization, and Collaboration
Project Motivation

- Teaching the “gateway” course since 2005
- Diversity of the ~40 ISchools
  - Disciplinary roots
  - Domain emphasis
  - Degrees offered
  - Student populations
  - Typical employers of graduates

The Missing Intersection

- The ISchools celebrate their disciplinary diversity, and are all “interested in the relationship between information, people and technology” but what do they really have in common?
- This diversity is inevitable and has some benefit, but has significant downsides
The “Gateway” Course Dilemma

What to teach? The Intersection is often described in terms of “Tradeoffs” but is not well defined...

The Tradeoff
Is there a more holistic framework for these topics that enables a coherent and common gateway course?

... wouldn’t such a framework define the intellectual intersection of the Ischools?

We Organize…

- Things
- Information
- Information about Things
- Information about Information about {Things, Information}

...
We Organize…

- Libraries, museums, business information systems, scientific data… and other institutional resource collections
- Different types of documents – from narrative to transactional – which have characteristic content, structures, and presentations
- Personal information and artifacts of all kinds in our kitchens, closets, personal computers, smartphones…
Museum

Louvre, Paris – with Mona

Photos by Bob Glushko

Archive
The Web & Digital Libraries

Retail Store
Web Retail Store

Real-Time Information About Information
We can emphasize how all of these domains and types of collections differ… or we can emphasize what they have in common.

They are all “Organizing Systems”

- A collection of resources
- Intentionally arranged
- To enable some set of interactions
RESOURCES are “anything of value that can support goal-oriented activity”

A COLLECTION is a group of resources that have been selected for some purpose
The Organizing System [2]

- INTENTIONAL ARRANGEMENT captures the idea that the system requires explicit or implicit acts of organization by AGENTS – human or computational ones
- These arrangements follow or embody one or more ORGANIZING PRINCIPLES

Organizing Principles & The Three-Tier Model

<table>
<thead>
<tr>
<th>In Software Architecture</th>
<th>In the Library Organizing System</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Interface / Presentation Tier</td>
<td>Library Services</td>
</tr>
<tr>
<td>Business Logic / Application Tier</td>
<td>Classification</td>
</tr>
<tr>
<td>Storage / Data Tier</td>
<td>The Books</td>
</tr>
</tbody>
</table>
Principles don’t Specify Implementation: “Organize Spices Alphabetically”

The Organizing System [3]

- INTERACTIONS include any activity, function, or service supported by or enabled with respect to the resources in a collection or with respect the collection as a whole

- Interactions can include access, reuse, copying, transforming, translating, comparing, combining… anything that a person or process can do with the resources…
The Activities of Organizing Systems

- We can identify four activities in the lifecycle of every organizing system:
  - Selecting resources
  - Organizing resources
  - Supporting resource-based interactions and services
  - Maintaining resources
For Your Closet Organizing System...

- **Selecting:** Should I hang up my sweaters in the closet or put them in a drawer?
- **Organizing:** Should I sort my shirts by color, sleeve type, or season?
- **Supporting Interactions:** Do I need separate places for laundry or dry cleaning?
- **Maintaining:** Should I toss out my clothes based only on stains and tears, based on how long I’ve owned them, or based on whether I’m tired of them?
Stop and Reflect…

- Does the definition of “Organizing System” apply to all of the types of collections, documents, and information services we’ve seen?

- What are the consequences of using abstract terms like “resource,” “interaction,” and “maintenance” instead of more specific and domain-specific terms?

  - Collection development, Appraisal vs. Selecting
  - Cataloguing, Indexing vs. Organizing
  - Curation, governance vs. Maintenance …

Resources…

In the Library?  In the Zoo?

On the Web:
Universal Resource Identifiers (URIs)
Organizing Organizing Systems [1]

• We can classify organizing systems by:
  • resource type
  • dominant purpose
  • creator
  • size of intended user community
  • or many other ways

Categorizing by Resource Type

- Collections of Books (Libraries)
- Collections of Art (Museum)
- Collections of Documents (Archive)
- Collections of Data (Repository)
- Collections of Spices (Pantry) …
The “Document Type Spectrum”

from “Document Engineering,”
R. Glushko & T. McGrath, MIT Press 2005

Narrative Document Type

CALL me Ishmael. Some years ago—never mind how long precisely—having little or no money in my purse, and nothing particular to interest me on shore, I thought I would sail about a little and see the watery part of the world. It is a way I have of driving off the spleen, and regulating the circulation. Whenever I find myself growing grim about the mouth; whenever it is a damp, drizzly November in my soul; whenever I find myself involuntarily pausing before a coffin warehouses, and bringing up the rear of every funeral I meet; and especially whenever my hypos get such an upper hand of me, that it requires a strong moral principle to prevent me from deliberately stepping into the street, and methodically knocking people’s hats off—then, I account it high time to get to sea as soon as I can. This is my substitute for pistol and ball. With a philosophical flourish Cato throws himself upon his sword; I quietly take to the ship. There is nothing surprising in this. If they but knew it, almost all men in their degree, some time or other, cherish very nearly the same feelings towards the ocean with me.
# Semi-Structured Document Type

Industrial Or Light Weight Bags On A Roll

<table>
<thead>
<tr>
<th>Bags Per Roll</th>
<th>Part No.</th>
<th>Price</th>
<th>Part No.</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 x 6</td>
<td>1000</td>
<td>25.30</td>
<td>8404LU</td>
<td>45.50</td>
</tr>
<tr>
<td>6 x 9</td>
<td>1000</td>
<td>30.10</td>
<td>8405LU</td>
<td>50.00</td>
</tr>
<tr>
<td>8 x 10</td>
<td>1000</td>
<td>47.10</td>
<td>8406LU</td>
<td>70.60</td>
</tr>
<tr>
<td>10 x 19</td>
<td>1000</td>
<td>66.20</td>
<td>8407LU</td>
<td>132.30</td>
</tr>
</tbody>
</table>

Discount per percent: Less 5% on 12 rolls; 10% on 24 rolls or more.

Lightweight Bags On A Roll

<table>
<thead>
<tr>
<th>Bags Per Roll</th>
<th>Part No.</th>
<th>Price Per Carton of 2 Rolls</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-11</td>
<td>1050</td>
<td>52.80</td>
</tr>
<tr>
<td>12-23</td>
<td>1050</td>
<td>56.16</td>
</tr>
<tr>
<td>24+</td>
<td>1050</td>
<td>44.88</td>
</tr>
</tbody>
</table>

Discount per percent: Less 5% on 12 rolls; 10% on 24 rolls or more.

# Transactional Document Type

**Energy Statement**

<table>
<thead>
<tr>
<th>Account Number</th>
<th>Bill Date</th>
<th>Amount Due</th>
<th>Due Date</th>
<th>Amount Enclosed</th>
</tr>
</thead>
<tbody>
<tr>
<td>69991.36-495.185825</td>
<td>06/19/2009</td>
<td>$155.50</td>
<td>12/07/2009</td>
<td></td>
</tr>
</tbody>
</table>

**Account Summary**

- **Service**: Gas
- **Service Dates**: 06/19/2009 to 09/16/2009
- **Amount**: $37.40
- **Energy Consumption Tax**: 0.13
- **Gas PPP Surcharge**: 1.84
- **Utility User Tax**: 15.71

**TOTAL CURRENT CHARGES**: $155.50

**Previous Balance**: 157.82
**OIBD Payment - Thank You**: 157.82

**TOTAL AMOUNT DUE**: $155.50
**DUE DATE**: 10/07/2009
Categorizing by Purpose: Resource Preservation as Means vs. End

Organizing Systems

Memory Institutions
- Libraries
- Museums
- Archives

Business Info Systems
- Content Management
- CRM
- ERP

Organizing Organizing Systems [2]

• Many more categorizations:
  • Personal collections vs. institutional collections
  • Location of the user community
  • Technology used
  • …

• But these classifications overlap without clear boundaries or necessary and sufficient features
A Library?

2012 Catalog
Welcome to our catalog for the 2012 growing season. It features over 150 varieties of seed, all of which have been expertly selected for their ability to grow well in your region. Browse our catalog by navigating the categories on the left or by using our full search guide on the next page.
A Library?

What is a Library?

- Collection of resources
- Organized to enable “access” and “reuse”
- Curated for “public good”
- Conventional interaction is “circulation” – borrowing and return of resources...

- Seed library and Wikipedia share some of these properties… but not the last… you only return a resource if you have improved its quality
Google Books == Library?

**LIBRARY**
- Non-Profit Objective
- No Data Retention – for Privacy
- Few Restrictions on Uses

**BUSINESS**
- Profit Objective
- Data Retention for Personalization
- Many Restrictions on Uses

SELECTED RESOURCES
- ORGANIZED RESOURCES
- RESOURCE-BASED INTERACTIONS
- PRESERVED RESOURCES

**YES:** www.nytimes.com/2009/10/09/opinion/09brin.html?

**NO:** http://www.huffingtonpost.com/pamela-samuelson/google-books-is-not-a-lib_b_317518.html

Consequences of Dimensional Thinking

- Many types of resource collections have conventional characteristics that are deeply embedded in culture and language

- Using an established category to describe an organizing system reinforces these characteristics, even if we add qualifiers (“seed” library)

- … and marginalizes any atypical characteristics of the organizing system being categorized
Natural History Museum

Zoo == Animal Museum?
Animal Theme Park == Museum? == Zoo?

Wild Animal Park == Zoo?
Colonial Williamsburg == Theme Park? == Living History Museum?

Human Resource Organization

Figure 1a
Organizational Structure

[Diagram not transcribed]
Cemetery == Memory Institution?
== Human Resource Organization?

Natural History Museum
A “Design Space” or “Dimensional” Perspective

• In addition to using categories like Library or Museum or Business Information System, consider a specific organizing system as a point in a multidimensional design space and these categories as regions in that space...

• This treats the familiar categories as “design patterns” that embody typical configurations of design choices
Consequences of Dimensional Thinking

• Overcomes the bias and conservatism inherent in familiar categories

• Design patterns support multi-disciplinary work that cuts across familiar categories and applies knowledge about them to new domains

• Creates a design vocabulary for translating concepts and concerns from category and discipline-specific vocabularies

The 5 Dimensions of an Organizing System

• What Is Being Organized?
• Why Is It Being Organized?
• How Much Is It Being Organized?
• When Is It Being Organized?
• Who (or What) is Organizing It?
What Is Being Organized?

- Identifying the unit of analysis is a central problem in every intellectual or scientific discipline - and in every organizing system.

- Resources that are aggregates or composites of other resources, or that have internal structure, pose questions about the granularity of their "thingness".

How Many Things?

SAN FRANCISCO'S WORLD CHAMPIONS

1981
How Many Things?

---

I STUFFED a shirt or two into my old carpet-bag, tucked it under my arm, and started for Cape Horn and the Pacific. Quitting the good city of old Manhatto, I duly arrived in New Bedford. It was on a Saturday night in December. Much
How Many Things?

I STUFFED a shirt or two into my old carpet-bag, tucked it under my arm, and started for Cape Horn and the Pacific. Quitting the good city of old Manhatto, I duly arrived in New Bedford. It was on a Saturday night in December. Much...

Design Choices & Patterns for Resources

- Books
- Art
- Music
- Data
- etc...

Domain, Focus, Resources, Format, Agency

Primary, Description, Active, Passive
Resource Formats

- Resource
  - Physical
  - Digital
    - Digitization
    - “Born Digital”

Format Matters!

INFORMATION IQ

Explicitness of Content

Separation of Control and Presentation

XML, HTML with CSS
Resource Focus

• We often designate some resource as primary because it is the focus of our attention
• We often create other resources that are descriptions of or otherwise associated with the primary resource
• We call these “Description resources” (a more general term than “metadata”)
• Primary and description resources can be either physical or digital…. So there are 4 combinations.
“Augmented Reality” – Digital Descriptions for Physical Resources

- San Francisco Nearest Transit: Floating signs can direct you to the nearest bus line or subway station.
- Wikitude: Tags for Wikipedia entries and points of interest float on screen, telling you what's around.
- Layar: Layar lets you set a radius and filter your search by restaurants, bars, or other destinations.

Printed QR “Quick Response” Codes are Physical Descriptions of Digital Resources

THE DISCIPLINE OF ORGANIZING

We organize things, we organize information, we organize information about things, and we organize information about information. For even though "organizing" is a fundamental and ubiquitous challenge, when we complete these activities there are more apparent than these communications. We propose to study the organization task by the concept of an Organizing System, defined as an information management collection of resources and the transactions that support them. Organizing Systems involves a collection of resources, a choice of properties or attributes used to describe and manage resources, and data of representing interactions with resources. By compiling and containing these resources, the activities take place in different contexts and domains, we always try to produce a balanced set of Organizing Systems.

TABLE OF CONTENTS
1. Foundations of Organizing Systems
2. Activities in Organizing Systems
3. Resources in Organizing Systems
4. Resource Description and Retrieval
5. Developing Relationships and Interactions
6. Organizing and Describing Resources
7. Describing and Interpreting Resources
8. Describing and Analyzing Resources
9. A Roadmap for Organizing Systems
Fantasy Football: One Person’s Description is another Person’s Resource

Description Resources
Aggregated Description Resources

Bibliographic Description Resources - 1920

1920's Catalog Card
Author Card, Title Card, Keyword catalog

Primary Resource
Author: Title
Publisher: ISBN:
Bibliographic Description Resources - 1960

1960's Machine Readable Cataloging
MARC Record

Primary Resource

1XX: Author
2XX: Title
3XX: Physical

Bibliographic Description Resources - 2000

University of California, Berkeley
School of Information
Resource Agency

• Passive or operand resources ("nouns") must be acted upon or interacted with to produce an effect

• Active or operant resources ("verbs") create effects or value on their own, sometimes by initiating interactions with operand resources

Smart Resources

Swiss Cows Send Texts to Announce They’re in Heat
NY Times 1 October 2012
Resources Over Time

Changes in Possession

Authenticity

Provenance

Persistence

Effectivity

2013
What Country Do You Live In?

• Kingdom of Yugoslavia (1929 - 1941)
• Independent State of Croatia (1941 - 1945)
• Federal People's Republic of Yugoslavia (1945 – 1963)
• State Union of Serbia and Montenegro (2003-2006)
• Serbia (2006-present)

Why Is It Being Organized?

• The essential purpose of an Organizing System is to "bring like things together and differentiating among them"

• But there are always more precise requirements and constraints to satisfy and more specific kinds of interactions to support
Interactions – The Why of Organizing Systems

- Some interactions can be enabled with any type of resource, while others are tied to resource types
- Interaction can be direct, mediated or indirect, or limited to interactions with resource copies or descriptions

Some Interactions with Resources

- **Search** (Boolean, Keyword, Structure)
- **Browse** (Faceted, Social)
- **Display** (Listen, Read, Watch, Perform)
- **Derivative** (Edit, Annotate, Augment, Mash Up, Transform)
- **Copy** (Share, Reproduce)
- **Sell** (Exchange, Transact, Resell)
- **Protect** (Preserve, Encrypt, Secure)
- **Destroy**
- **Organize** (Bookmark, Tag, Categorize, Move, Label)

Physical manipulation is an intrinsic interaction with collections of physical resources.

Physical manipulation and interpersonal contact might be required to interact with information resources in physical form like the printed books in libraries.

With digital resources, neither physical manipulation nor interpersonal contact is required for interactions, and the essence of the interaction is information exchange or symbolic manipulation of the information contained in the resource.

So we often interact with physical resources through associated digital resources.
Principles of Organization

• The organization of physical resources is strongly influenced by their material manifestation, and this organization often persists when digitized resources are organized.

• The simplest organizing principle is co-location.

• Almost any property of a resource might be used as a basis for its arrangement, and multiple properties are often used simultaneously.

• Properties of the collection as a whole can also be used in organizing principles.
Organizing System: Home Kitchen

Principles of Kitchen Organization

• **Intrinsic static properties**: If you store your pots, frying pans, and baking pans in different cabinets and nest each set by size.

• **Extrinsic static properties**: A spice rack with the spices arranged in alphabetical order.

• **Intrinsic dynamic properties** if you arrange your milk and other perishable goods by expiration date, a “useful life remaining” property that decreases to zero as the expiration date approaches.

• **Extrinsic dynamic properties** if you put the most frequently used condiments or spices in the front of a refrigerator or pantry shelf.
Principles of Document Organization

- **Intrinsic static properties**: Author, date published, words in the text
- **Extrinsic static properties**: ISBN, LOC Classifications
- **Intrinsic dynamic properties**: Effectivity (e.g., laws and regulations)
- **Extrinsic dynamic properties**: Links/citations to and from other documents

An Individual’s Book Classification System

Photo by Brendan Curran
A DJ Organizes His Record Collection

*This is the Hip-hop and Dancehall (Jamaican) section of my record collection.*

*When I return hip-hop records to the shelf, they just go in the front of that particular pile, with the net result that stuff that gets played less drifts to the back.*

Photo by Matt Earp aka Kid kameleon

---

How Much Is It Being Organized?

- Not every resource needs the same amount of organization
- Not everything is equally describable
- A controlled vocabulary can yield more consistent organization
- The scope and size of a collection shapes how much it needs to be organize
- Are you organizing the resources you have, or do you need to create an organizing system that can apply to resources that you have not yet collected?
When Is It Being Organized?

- When the resource is created
- When it is added to some collection
- Just in time
- Never
- All the time - continuous or incremental

“Just in Case” Organization
Postponing Organization

Who or What Is Organizing?

• Authors or creators
• Professional organizers
• Users “in the wild”
• Users "in institutional contexts“
• Automated or computerized processes
Summary

- The concept of Organizing System unifies a vast body of design and analysis practice from many disciplines.
- Thinking in terms of design dimensions overcomes the limitations and inertia of the traditional categories.
- It is a generative, forward-looking approach that encourages and accommodates innovation while preserving conventional theory and practice as design patterns.
- It enables intelligent conversations between people who didn’t have much common language before.
Project Collaboration

- This book began as the lecture notes from my Berkeley course, but I soon realized that it was more ambitious than I could do on my own.
- So I enlisted numerous collaborators, and the book is now the product of countless discussions with students and faculty colleagues at Berkeley and other schools.
- Manuscript is being used at Berkeley (2x), UNC (2x), Humboldt, and Kentucky; other schools will soon try it (Michigan, Illinois, Haifa, St Louis).

What We Are Publishing

- “The Discipline of Organizing” will be published by MIT Press in early 2013 (R. Glushko, editor) simultaneously in several different formats:
  - As a traditional printed book
  - As HTML files, freely available for anyone to use
  - In one or more ebook formats
    - Current generation – kindle etc
    - Next generation – embodies the transdisciplinary character of TDO.
Customization

- We have devised an approach that lets the book focus on core ideas while enabling it to satisfy different topical emphases ("flavors" or "channels")
- Solution: "Lean text" + tagged "endnotes"
  - Printed book has [LIS], [Computing], [CogSci], [Business], [Law], and [Citation]
  - E-Books might have [Highlight], [Annotation],[SelfStudy], [Interactive], and other channels

Example of Tagged Notes: Main Text

The Organizing System for a small collection can sometimes use only the minimal or default organizing principle of "co-location" – putting all the resources in the same container, on the same shelf, or in the same email inbox. If you don't cook much and have only a small number of spices in your kitchen, you don't need to alphabetize them because it is easy to find the one you want.
For collections of non-trivial size the choice of searching or sorting algorithm is a critical design decision because they differ greatly in the time they take to complete and the storage space they require. For example, if the collection is arranged in an unorganized or random manner (as a “pile”) and every resource must be examined, the time to find a particular item increases linearly with the collection size. If the collection is maintained in an ordered manner, a binary search algorithm can locate any item in a time proportional to the logarithm of the number of items. Analysis of algorithms is a fundamental topic in computer science; a popular textbook is “Introduction to Algorithms” by Thomas Cormen et al (2009).

---

Some digital formats support interactions that are qualitatively different and more powerful than those possible with physical resources. However, processing such interactions can be challenging, and the success of a digital resource depends on the design of the system. The physical form of the resource is a critical factor, as it can influence how the resource is perceived and used. For example, a physical book can provide a sense of ownership and a physical experience that is not possible with a digital version. In contrast, a digital copy can be easily shared and accessed from anywhere, which can be a significant advantage in the context of collaboration and dissemination.

Moreover, digital resources are typically more accessible and easier to manage than their physical counterparts. They can be easily searched, indexed, and organized, which can greatly improve the efficiency of resource discovery and use. Digital resources also offer a range of additional features, such as the ability to interlink with other resources, to provide annotations and comments, and to support interactive elements, which can enhance the user experience and increase engagement.

---

The Discipline of Organizing

Design Prototype (Foeckler & Nahman)
Customization with Channel Insertion

Some digital formats support interactions that are qualitatively different and more powerful than those available with physical documents. The growth of digital document manipulation technology, for instance, has led to the development of tools that enable users to interact with digital documents in new ways. For example, Google's Knowledge Panel tool can automatically highlight and expand definitions of entities within a document. This feature allows users to gain deeper insights into the content of the document without having to navigate through the entire document manually.

3.2.3 Resource Agency

Agency for resource sharing is a fundamental aspect of the digital environment. In this section, we explore various methods for agency in the digital world.

Reader Highlighting

Some digital formats support interactions that are qualitatively different and more powerful than those available with physical documents. The growth of digital document manipulation technology, for instance, has led to the development of tools that enable users to interact with digital documents in new ways. For example, Google's Knowledge Panel tool can automatically highlight and expand definitions of entities within a document. This feature allows users to gain deeper insights into the content of the document without having to navigate through the entire document manually.

3.2.3 Resource Agency

Agency for resource sharing is a fundamental aspect of the digital environment. In this section, we explore various methods for agency in the digital world.
What Does the Professor Highlight?

SCHOOL OF INFORMATION
UNIVERSITY OF CALIFORNIA, BERKELEY

Self-Study Channel
Collaboration and Curation Mechanisms

- [DisciplineOfOrganizing.org](http://tdo.berkeley.edu)
- Dropbox for sharing lecture notes, assignments, exams, and other teaching materials
- **TDO video channel** for sharing recorded lectures
- Bi-monthly conference calls
- Work underway to design collaborative publishing system so that TDO can be maintained by the ISchools

Acknowledgements

- Thanks to 3 faculty collaborators, 11 ISchool master’s students (2010-2012), 2 PhD students from UCLA who are co-authors of the book
- Thanks to the dozen people who have helped with editing, bibliography, graphics, and eBook design
- Thanks to 12 teaching assistants and more than 100 students who have put up with my evolving understanding and systematization of TDO
- Thanks to Dean Saxenian and my faculty colleagues at Berkeley for providing a laboratory for intellectual experimentation
Comments, questions?

glushko@berkeley.edu