Conceptualizing Institutional Repositories:
Using Co-Discovery to Uncover Mental Models

Soo Young Rieh, Ji Yeon Yang, Elizabeth Yakel, Karen Markey
School of Information, University of Michigan
304 West Hall, 1085 South University Avenue, Ann Arbor, MI 48109-1107 USA
+1 734.647.8040
rieh@umich.edu, jiyeon@umich.edu, yakel@umich.edu, ylime@umich.edu

ABSTRACT

This study investigates how people construct mental models of new information systems with which they have limited experience. Six different institutional repositories were used as the experimental systems for this lab-based co-discovery experimental study. Sixty subjects (30 pairs) were asked to complete search tasks based on a simulated work situations using an institutional repository. Subsequently, subjects were instructed to visually depict how they thought the institutional repository worked and then explain this to their partner. Our findings are based on these drawings, descriptors written on drawings, and audio-recordings of explanations and conversations. The results reveal that most of the subjects constructed mental models focusing on system operations and the design of the user interface. Few highlighted the interactivity between the system and the end user or presented a global-view of the system to show how it related to other search engines or databases. We found that the co-discovery method provides a viable research design to elicit people's mental model construction. The implications of the results for interactive information retrieval community and institutional repository community are discussed in terms of research design, search behavior, and user instruction.

Categories and Subject Descriptors

H.3.5.Online Information Services, Web-based services; H.3.7. Digital Libraries, User issues

General Terms

Human Factors

Keywords

Mental model of an information retrieval system; Co-discovery method; Institutional repositories

1. INTRODUCTION

People's understandings of "how an IR system works" can provide useful insights to explain their information searching behaviors. For example, the ways people conceive of a system are directly related to the patterns and strategies to be used during

information searching [1]. Examining conceptions, particularly prior knowledge and beliefs about the purpose and function of a system, that people develop to make sense of IR systems also informs us about how people learn to use them [2]. Given this research on search behaviors and the importance of how people envision systems to the search process, we designed a study to better understand how subjects conceptualize a new type of IR system. Our findings highlight which core functions and features people immediately recognize from a search interaction and what salient elements they ignore in an initial encounter with a new system.

We selected a mental model framework to guide our research. Mental models are users' conceptual or operational representations of a system [3]. By uncovering or eliciting a mental model, we can learn whether a user can predict the operation of a system and, eventually, build an image of the system to better guide future tasks [3]. Mental models describe users' knowledge about how a system works as well as its components, processes, structural relationships between components, and influences of various events in the system [4, 5]. Therefore, mental models inform our understanding of users' behaviors, such as errors or choices [6].

An earlier study using mental models of information retrieval systems was conducted by Borgman [7]. She examined how undergraduate students explained the operation of online library catalogs. She found that naïve subjects, who had little or no prior computing experience, performed better when trained using a mental model of the system than those who were given only procedural instructions; however, these findings only applied to complex, problem-solving search tasks. More recently, Hendry and Efthimiadis [8] prompted 232 information science undergraduate and graduate students to draw sketches of how they thought search engines functioned. Their findings reveal that students produced sketches describing only a few aspects of the search engines' functionality. The sketches also demonstrate the diverse approaches students took to express their understandings of search engines – algorithms, iconographic depictions, and metaphoric languages. Zhang [1] also investigated the relationship between undergraduate students' mental models and their searching behaviors. Her results show that students who perceive search engines as the center of the Web (process view) perform the best, but have more difficulty with the tasks and are least satisfied with their performance. On the other hand, students who view the Web as a collection of computers, services, modems, and CPUs (technical view) have less difficulty with the search tasks and are more satisfied with their performance.

The present study investigates undergraduate students' construction of mental models for institutional repositories. Since 2002, institutional repositories have been deployed by colleges

and universities to collect, preserve, and provide access to digital content produced within university community [9]. While academic librarians have been active initiating institutional repositories [10], the conception of what institutional repositories are is still new to most people and even to its end-users [11]. Research has shown that students' conceptualizations of new systems are influenced by their current understandings of the dominant IR systems, such as Google [12, 13]. This study builds on these findings by focusing on investigating mental models of a new and unfamiliar information system, such as an institutional repository, based on the user's first experience of the system. Unlike most empirical studies on mental models of the Internet or search engines [1, 2, 8], this study aims to analyze the conceptual depictions and specific descriptors undergraduate students used in order to represent their understandings of institutional repositories.

To better understand people's knowledge about how an institutional repository works, we use a co-discovery method in our experimental laboratory study. Using the co-discovery method, paired subjects are encouraged to help each other accomplish a common goal, to explain what they are doing, and articulate why they are doing it while they are working on the tasks. Compared to a single subject think-aloud protocol, the conversational nature of the technique makes it more natural for study subjects to verbalize their thoughts during the experiment [14, 15]. This technique has been used before to help subjects articulate their thought processes [16]. This research design provides an opportunity for each pair to engage in extensive discussions about how the system works.

This article addresses the following research questions.

1. How do people visually represent the operation of an institutional repository?
2. To what extent do people understand the nature and scope of an institutional repository?
3. Is the co-discovery method viable for eliciting the construction of people's mental models?

2. METHOD

2.1 System Selection
Since 2002, academic institutions have offered institutional repository systems and services to their learning communities [10]. Institutional repositories are implemented using a variety of open-source and proprietary digital library content management systems, such as DSpace, EPrints, bepress, Fedora, and Content DM to enable colleges and universities to publish their digital content. While academic library managers and staff see institutional repositories as a unique opportunity to make a set of new online services available to their university communities and increase access to digital content [9], most end-users of institutional repositories are uncertain about the scope and purposes of these systems [11]. Interviewees in St. Jean et al.'s study expressed substantial confusion about how institutional repositories differed from other library databases [11]. To better understand how people conceptualize institutional repositories, we conducted an experimental study to investigate users' mental models of these systems.

Unlike online library systems or Web search engines, institutional repositories contain unique local content and may have locally customized interfaces and functionalities. Therefore, we identified diverse institutional repositories as experimental systems. The six institutional repositories selected were: Deep Blue (University of Michigan), Knowledge Bank (Ohio State University), eScholarship (University of California/California Digital Library), DSpace (MIT), DSpace (University of Delaware), and iDEA (Drexel University). All of the institutional repositories run on DSpace except eScholarship which uses bepress. Various factors were considered in selecting these repositories. DSpace@MIT was chosen because MIT collaborated with Hewlett Packard to develop DSpace in 2002 and it is one of most mature institutional repositories. eScholarship was also included because of the large number of participating research units (279 units as of February 2010) and its emphasis on open access. The other four systems were mainly selected because each system enabled us to conduct a particular retrieval task that ranked in the top 10 results on a Google search results page and retrieved different genres of material (e.g., research reports, archival and manuscript material).

2.2 Subjects
We collected data from 60 undergraduate students who were 18 year or older at a Midwest university in the United States in March and April of 2009. Subjects were recruited through flyers posted in various places on campus. The flyer did not include a phrase of "institutional repository" or "mental model." It simply said to "bring a friend, search the web together, tell us what you find, and earn $30 each." Friends came to the lab at the same time, and worked together in experimental sessions. We wanted existing friends because we thought this would be more likely to generate active conversations.

Of the 60 subjects, 36 subjects were female (60%) and 24 were male (40%) students. Their ages ranged from 18 to 25, and the average age was 20.1 years old. Ten freshmen (16.7%), 20 sophomores (33.3%), 12 juniors (20%), and 18 seniors (30%) from various academic backgrounds across humanities, social science, science, and engineering participated. None of the subjects had previous experience using an institutional repository. Subjects reported that they did not frequently use online catalogs or databases. Out of 60 subjects, 14 (23.3%) said that they used an online catalog at least once a week. Twelve subjects (20%) responded that they used the online catalog only once a month while 14 (23.3%) used it several times a year. The frequency of online database use was even lower. Only 16 of subjects (26%) said they used an online database at least once a month. There were 14 subjects (23.3%) who had never used the online databases available through the library. On the other hand, all of subjects responded that they were confident in finding information on the Web. Thirty two subjects (53.3%) responded that they were "very confident" and the remaining 28 subjects (46.6%) said that they were "somewhat confident."

2.3 Tasks Used in Experiments
We developed tasks that simulated a work situation and provided some context in a scenario for the given task [17]. Given the fact that subjects were undergraduate students, we developed simulated information needs that were familiar to them – researching a term paper (Task 1), creating a presentation (task 2 and Task 5), pursuing academic interests (Task 3), locating class readings (Task 4), and helping a friend (Task 6). In terms of the search itself, we wanted subjects to complete tasks that required them to find a set of documents, a series of research reports, or archival materials which matched their simulated information need to a specific institutional repository. Through a series of pre-tests, we developed tasks in such a way that subjects would likely

find the results for each task in a targeted institutional repository – Task 1: University of Michigan Deep Blue, Task 2: Knowledge Bank, Task 3: eScholarship, Task 4: DSpace@MIT, Task 5: DSpace@ University of Delaware, and Task 6: iDEA.

The six search tasks used in the study were as follows.

- Task 1: You are interested in how computer simulation is applied in transportation related studies. You talked with your instructor about the possibility of choosing this topic for your term paper. Your instructor suggested that you could start with the "Transportation Research Institute (UMTRI)" collection. He also said that in order to write a good paper, you should try to find multiple articles using computer simulation in transportation related studies, such as predicting the braking performance or steering response.

- Task 2: Your history professor told you that there are very good photos on polar exploration taken by Rear Admiral Richard E. Byrd. He said that the collection called Byrd Expedition Photo Albums contains a lot of good images that you could use for your presentation. Please try to locate this collection and find a few photos of expedition members, dogs, or any other interesting images that you want to use.

- Task 3: You want to know more about energy efficiency. As your friend studies at University of California Energy Institute, you want to search their collection to find multiple articles on energy efficiency.

- Task 4: You major in Electrical Engineering and you hear that there is a collection called "MIT OpenCourseWare - Archived Content". You want to find some syllabus or readings for Electrical Engineering field there.

- Task 5: You are planning a presentation on public education in Delaware. You know that Delaware Education Research and Development Center may have something helpful on your topic. You want to explore their collection to find multiple sources for your presentation.

- Task 6: Your friend asked you to find some recent studies on art therapy. You think it might be a good idea to search Drexel Theses and Dissertations collection first. What can you find there for your friend?

2.4 Procedures
The actual sequence of procedures for the experimental study follows. Once a pair of subjects arrived, each person was asked to fill out and sign the consent form. The general instructions provided in the informed consent were: "You and your friend will be given two topics to search in two different online systems. While you search, we will ask you to tell your friend what you are doing and why." Two search tasks were assigned to a pair of subjects using the Latin Squares method so that every task was used by ten different pairs and that the presentation order of the institutional repositories was varied. One subject used the mouse and the keyboard for the first task; the other for the second. Each

subject was asked to begin the task with Google. Google was selected as a starting point because users often begin searches with a search engine and institutional repositories advertise that their materials are indexed by Google [18]. Subjects worked on each task for 10 minutes. Once a subject completed the session on Google, the researcher asked a few questions about his or her reasons for selecting particular links on the search results page. After the Google portion of the task was completed, the institutional repository which contained information pertinent to the search task was introduced. The subject was then asked to complete the same task using a particular institutional repository for another 10 minutes. After that, the subject was instructed to draw how he thought the institutional repository worked and then to explain the drawing to the other subject. The other subject then had a chance to make changes to the drawing marking with a red pen based on his or her own observations. The procedures – (1) searching on Google; (2) searching in an institutional repository; (3) drawing how the institutional repository works; and (4) making changes to the drawing – were then repeated for the second searching task, with the other subject now taking the lead role. At the end of each session of the experiment, an exit interview was conducted in order to investigate the subjects' perceptions of an institutional repository and evaluations of the searching experience. We also collected the subjects' demographic background information at the end of the session. Searching activity and an audio of the conversation between subjects were recorded using Camtasia software.

2.5 Data Analysis
In this paper, our research questions concern the development of the mental models. Therefore, our analyses focus on the drawings, descriptors written on drawings, and audio-recordings of subjects' explanations of their mental models and the discussions between the two subjects. We examine subjects' understandings of institutional repositories to identify the extent to which they were able to ascertain the features, functions, and content of an institutional repository quickly and what misconceptions arose. For the analysis of drawings, we identified four distinct approaches people have taken to explain how systems operate: processing, global view, interface, and interactivity (Table 1). Although our initial analysis was guided by Zhang's [1] four views – technical, functional, process, and connection –in the end, we found that Zhang's typology did not match our emerging findings. This may be because our subjects drew a system with which they had minimal experience whereas Zhang's subjects drew mental models about the Web, with which they were familiar. In the process of drawing mental models, subjects included a variety of descriptors to label events, interaction, system, and collections. A content analysis of subjects' descriptors on the drawings was conducted to identify general concepts which emerged from their initial encounters with institutional repositories. Lastly, we analyzed transcriptions of conversations between the two subjects during our co-discovery experiments to characterize patterns of negotiation, compromises, and disagreement about their mental models.

Table 1. Categories for Data Analysis

| Category	Subcategory	Definition
Approaches to explaining the operation	Processing model	Subjects illustrate an institutional repository as system operations, infrastructure, or an algorithm that enables to execute the storage and retrieval of information
	Global-view model	Subjects view an institutional repository associated with other information systems such as the internet, search engines, websites, and library databases
	Interface model	Subjects depict an institutional repository in terms of design of user interface
	Interactivity model	Subjects represent human activities in an institutional repository
Understanding	Full understanding	Subjects fully understand the nature and scope of an institutional repository
	Partial understanding	Subjects partially understand an institutional repository by identifying the name of the institution, institutional units, or any unique attributes to an institutional repository
	Misunderstanding	Subjects show misunderstanding of an institutional repository based on inaccurate representations of system attributes or concepts
	Not understanding	Subjects do not understand an institutional repository at all evidenced by very few descriptors or incomplete drawings
Manifestations of the concept	Interaction	Descriptors representing interaction between a user and an institutional repositories such as search, functionalities, results, and criteria
	System dimensions	Descriptors representing system dimensions of institutional repositories such as repository collections and information system infrastructure
	Institutional units	Descriptors representing institutional units or brands of institutional repositories
Co-discovery patterns		Constructive interaction between the paired subjects about their mental models through either revision of drawings or conversation

3. RESULTS

3.1 Approaches to Explaining Operation of Institutional Repositories

We categorized subjects' drawings into four models: Processing, Global-view, Interface, and Interactivity. Model classification was mutually exclusive. The results show that the Processing model was depicted most frequently (n=38, 63%), followed by the Interface (n=11, 18%), Interactivity (n=6, 10%), and Global-view (n=5, 8%) models.

Processing models illustrate the multiple components of a system. More than half of our subjects viewed institutional repositories in terms of system operations, infrastructure, or an algorithm enabling the retrieval and storage of information. For instance, input devices (e.g., search boxes or browsing choices), output search results, and matching algorithms often appeared in the drawings. Search functions were often shown in relation to a server, multiple computer terminals, and databases. The components of institutional repositories were represented in terms of collections, institutional units, or the branding of the institutional repositories. The relationships between or among the components of repositories were often depicted using metaphors, such as trees, flow charts, stars, or nests.

A majority of the students' processing models represented general features of search systems rather than focusing on the unique characteristics in each institutional repository. For instance, subject S2B's mental model of eScholarship contained only three components: "catalog," "database," and "related articles." He simply connected these three components in his drawing. Subject S5A's mental model of Deep Blue also depicted a barebones process from "key word search" to "articles relevant" in terms of

"key words, titles, organizations sponsoring paper, author". She also included a note, "entire data base," in her drawing. A similar example can be found in subject S6A's drawing of Knowledge Bank. Using the tree metaphor, she started her processing model with "home" leading to "search" then to a set of "results" that took another round of "search" and "results" until "results" reached "image."
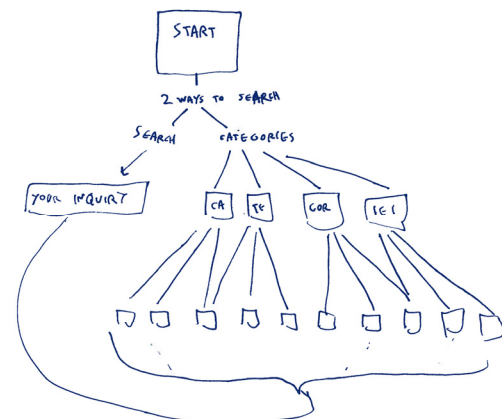
Figure 1. Processing model (Subject S24B, DSpace@MIT)

Figure 1 shows a typical example of the processing model. In this drawing, subject S24B, who used DSpace@MIT, first drew a "start" point. He illustrated "2 ways to search" offering two paths of "search" and browsing "categories." He then connected the process of search to "your inquiry" while depicting four types of collections. The collections were then distributed into numerous

individual documents. The drawing then illustrated the matching process between terms entered through "your inquiry" and some sort of representation of documents. This drawing indicates that subject S24B had gained enough familiarity with DSpace@MIT to form a rudimentary conceptualization of the system. Although DSpace presents two ways of accessing the collection – "Search" and "Communities," subject S24B used a more generic term, e.g. "categories," in his drawing. He also did not include descriptors, such as MIT, DSpace, or an institutional repository.

In the *global-view model,* subjects' depicted institutional repositories as one system among many, including the Internet, search engines, websites, and/or courseware. Subjects made attempts to better understand an institutional repository by positioning it alongside other systems with which they had greater familiarity. However, due to their limited knowledge of institutional repositories, their conceptions were often not accurate. In fact, of the 5 drawings categorized as global-view models, three misunderstood the nature and scope of an institutional repository and the other two subjects only had a partial understanding. For instance, subject S2A drew a model where the documents were retrieved from Deep Blue through the "internet." His drawings presented that "internet" as mediating between the collection and Deep Blue. He did not specify what types of the documents were contained in Deep Blue or any selection process. Another subject, S22A, who depicted both a search engine and DSpace at the University of Delaware, did not accurately understand the relationship between the search engine and DSpace. Subject S22A depicted documents from different sources, such as "The Delaware" and "The NYTimes", all connected to computer servers. The document eventually reached "DSpace" through a search engine. The metaphor he used for the search engine was a car, something that gets you where you want to go (See Figure 7).

In Figure 2, subject S3B illustrated DSpace's association with another system on MIT campus, Open Course Ware. Her drawing indicated that she believed that DSpace was for "research" whereas Open Course Ware was for "current courses." She told her partner that she could directly access MIT's Open Course Ware without using DSpace to retrieve course information. Both her drawing and the conversation with her partner indicated that she was confused about the relationship between these two systems.
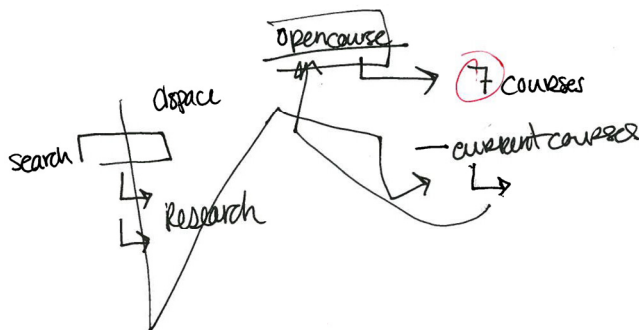

Figure 2. Global model (Subject S3B, DSpace@MIT)

Interface models focus on the design features of the user interface. Subjects sketched what they saw on the computer screen, including wireframes of web pages, search boxes, search results lists, links, tabs, and logos. Out of eleven subjects who constructed Interface models, five included multiple wireframes to show a sequence of search processes while six represented the interface with a single wireframe. Subjects who drew one screen tended to represent the search results page rather than the homepage of the institutional repositories.

Figure 3 shows one example of the *Interface model.* Subject S1B drew a series of wireframes to explain the operation of Knowledge Bank. The first screen contains a "search engine" box under the "Knowledge Bank" brand. Her next screen illustrates how to search within a particular "collection." She then sketches the search results interface in which three photos and descriptions are listed. In the last wireframe, she places the "photo" for which she was looking. This particular example shows that subject S1B has constructed a mental model centered on the user interface and navigation through the site. Even though she did not recall the details of Knowledge Bank's user interface, she remembered the navigational elements of searching within a particular collection.
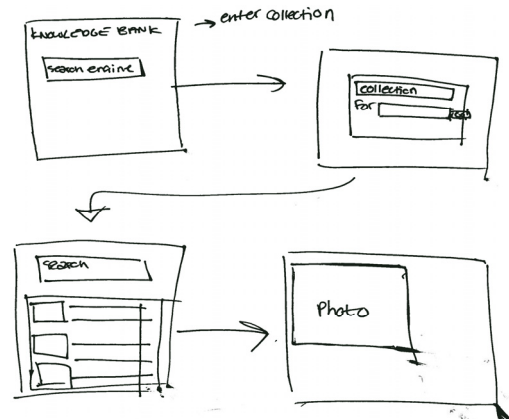

Figure 3. Interface model (Subject S1B, Knowledge Bank)

Interactivity models focus on human activities within a system. Subjects who depicted institutional repositories in this way included such actions as type, click, search, browse, read, view, and revise. Students possessing this mental model also represented system activities, such as bring up a list, pull, match, display, and notify. Overall, subjects most frequently drew two interactive aspects: search and browse. Interestingly, the only depictions of humans or human actions occurred in representations of this model.
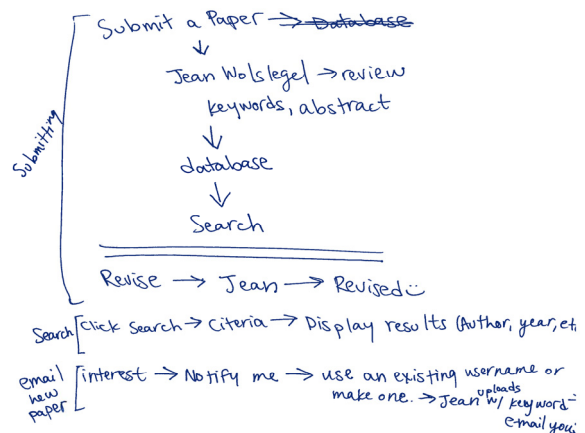

Figure 4. Interactivity model (Subject S15A, eScholarship)

In Figure 4, subject S15A rendered his view of the interactive process of eScholarship at University of California. Three stages of search are shown: "submitting," "search," and "email new paper." Within each stage, he conceptualized a detailed interactive process among a "paper," a "database," and a user's "interest." He also explained the nature of interactivity by including both system (display, notify me) and human activities (search, revise, click).

Table 2 provides an overview of which mental models were associated with which institutional repositories. Across the six different institutional repositories, the *Processing model* was dominant; 38 subjects depicted this approach. The *Interface model* was the next frequent; 11 subjects, fairly evenly across all the institutional repositories, drew diagrams which emphasized elements of an interface. Five subjects in total; 2 each from DSpace (MIT) and DSpace (Delaware) and one using Deep Blue (Michigan) developed *Global-view models*. No subjects using Knowledge Bank (Ohio State), eScholarship (California), or iDEA (Drexel) expressed a *Global-view model* or associated the institutional repository with other information systems. Only 6 subjects drew diagrams or annotated their representations with information to indicate the key *Interactivity model* feature: interaction between the system and the user.

**Table 2. Types of Models by Institutional Repositories**

Institutional Repository	System Types	Processing	Global-view	Interface	Interactivity	Total
Deep Blue – Michigan	DSpace	6	1	2	1	10
Knowledge Bank – Ohio State	DSpace	8	0	1	1	10
eScholarship Repository – California	bepress	7	0	2	1	10
DSpace @ MIT	DSpace	5	2	2	1	10
DSpace-Delaware	DSpace	5	2	3	0	10
iDEA: Drexel E-repository and Archives	DSpace	7	0	1	2	10
Total		38 63%	5 8%	11 18%	6 10%	60 100%

3.2 Understanding the Nature of Institutional Repositories

We used four categories to characterize subjects' level of understanding of the nature of institutional repositories: full understanding, partial understanding, misunderstanding, and not understanding. One way in which we measured understanding of the institutional repositories was to assess whether the subject understood that the institutional repository was associated with a specific institution (branding) and that it contained content from different units within that institution. There were only ten subjects (16.7%) out of 60 who seemed to fully understand the institutional nature and scope of institutional repositories. These subjects wrote

down either the institution name or academic/research units of the institution. These drawings also included specific attributes unique to an institution which went beyond depicting a general searching system or digital collections.

In Figure 5, subject S10B understood that iDEA was organized in terms of communities within the university, that each community represented a different domain area, and that multiple types of information resources, such as theses or dissertations, were contained in the system.
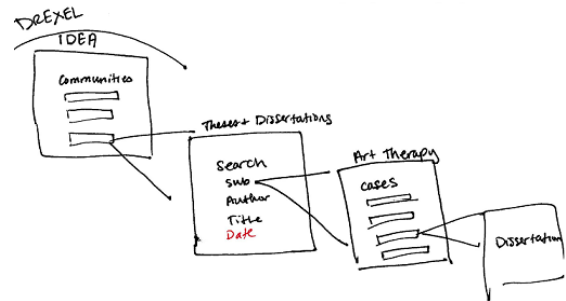
Figure 5. Full understanding (Subject S10B, iDEA)

Twenty six subjects (43.3%) had a partial understanding of an institutional repository. These subjects were not able to identify the name of the institution, institutional units, or any unique attributes related to the nature of institutional repository. They tended to construct their diagrams of an institutional repository as a generic information retrieval system possessing collection, interface, and search features.

Using a food processing metaphor going through a series of strainers, Subject S27A's drawing (Figure 6) depicts iDEA as an "archives" and illustrates how domain areas in an archival collection are narrowed down through filters and eventually presented on the "screen." His search was Task 6 which asked subjects to find recent studies on art therapy in Drexel's Theses and Dissertations collection. Although he viewed iDEA as a repository based on a filtering mechanism to retrieve information, there was no evidence that Subject S27A fully understood the nature or scope of the repository.
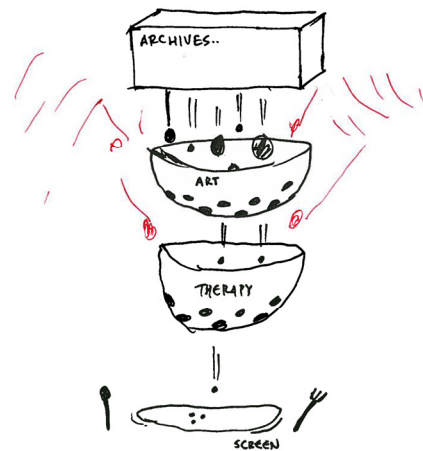
Figure 6. Partial understanding (Subject S27A, iDEA)

We identified nine cases (15%) in which institutional repositories were misrepresented by subjects. A common error occurred when subjects tried to conceptualize an institutional repository by

linking it to other information systems. Subjects then inaccurately specified the relationships between the institutional repository and the Internet or a search engine. Subject S2B viewed the institutional repository as a subsystem of the library's databases. Subject S3A thought that Deep Blue and other universities' repositories were an amalgamation of one large institutional repository. Subject S10B corrected subject S10A's original drawing saying that several different university libraries' collections contributed to Knowledge Bank. Subject S22A believed that the information from a variety of websites, including the New York Times, were included in the DSpace at the University of Delaware (Figure 7).
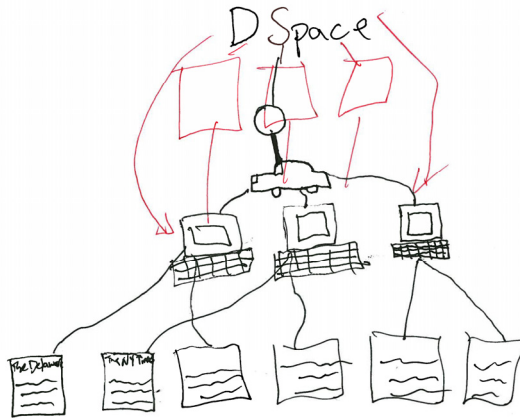
Figure 7. Misunderstanding (Subject S22A, DSpace at University Delaware)

On the other hand, 15 subjects (25%) did not understand institutional repositories at all. Their drawings were very simple and contained only one or two descriptors. In many cases, their drawings represented a search engine or the Internet generally, rather than the institutional repository. In other cases, subjects did not seem to have much idea how to represent an institutional repository because of their limited understanding, as seen in the case of subject S14B's drawing (Figure 8).
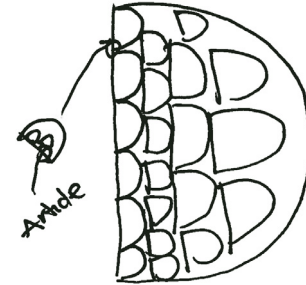
Figure 8. Not understanding (Subject S14B, DSpace at University of Delaware)

3.3   Content Analysis of Descriptors

From the drawings of the mental models, we identified 58 unique key terms out of 366 descriptors and categorized them into three major categories: Interaction, System dimensions, and Institutional units. We further categorized these into eleven sub-categories including search, functionalities, results, criteria, genre/form, metadata, collection, Web, system, institutions and academic units, and institutional repository brands (see Table 3).

Table 3. Descriptors Used in Mental Models

Major Categories	Sub-Categories		Unique Key Terms
Interaction (n=162)	Search (97)		Search (75); Keyword (16); Inquiry, term, query (6)
	Functionalities (29)		Email notification (6); Revise, redefine, repeat (4); Match (3); Browse (3); Bookmark (2); Code(d) (2); Options (2); Choose (1); Filter (1); Review (1); Scan (1); FAQ (1); Navigate (1); Read (1)
	Results (24)		Results (18); Links (3); Desired information (3)
	Criteria (12)		Relevance (7); Interest (3); Criteria (1); Credibility (1)
System dimensions (n=134)	Repository Collections (107)	Genre/ Form (47)	Articles, documents (21); Photos, image (8); Category (4); PDFs (4); Theses, dissertations (4); Journals (3); Research (2); Books (1)
		Metadata (42)	Topics, subjects (15); Author (10); Title (8); Abstract (4); Year, date (4); Catalog (1)
		Collection (18)	Collection (11); Albums (4); Archives, repository (3)
	Info-system Infrastructure (27)	Web (14)	Homepage (7); Search engine (4); Internet (1); Browser (1); Webpages (1)
		System (13)	Database (8); Open courseware (2); Screen (1); HTML (1); Server (1)
Institutional units (n=70)	Institutions and Academic Units (39)		Departments, branches, college, campus (9); University (8); Communities (7); Class, course (5); Institute, centers (5); Organizations (2); University libraries (2); Research unit (1)
	Brands (31)		DSpace (13); iDEA (8); Knowledge Bank (5); Deep Blue (3); eScholarship (2)

In terms of the major categories, a majority of the descriptors (162) represented elements associated with some type of interactions with institutional repositories, followed by the ones related to the system dimensions (134) and institutional units (70). Among the sub-categories, the terms that indicated some aspect of repository collections (107) were used the most, followed by search-related activities (97). Other popular types of terms included: genre/form (47), metadata (42), and institutions and academic units (39).

Subjects used the term "search" most frequently (75 times); more frequently than any other single term. Other commonly assigned terms were: articles/documents (21), results (18), keywords (16), topics/subjects (15), D-space (13), collection (11), and author (10). Doing a mashup of these terms we could say that a majority of the subjects perceived institutional repositories as "article keyword search systems." The high frequency of the term "search" indicates a prevailing view of this study's subjects that the institutional repositories are a type of an information retrieval system.

3.4 Co-Discovery by Paired Subjects

In the co-discovery method, subjects have the opportunity to comment or negotiate changes on their partner's drawings. Analysis of the drawings and conversations between paired subjects revealed that even when experiencing the same institutional repository together, two collaborators often took different approaches in constructing their mental models. Out of 30 paired subjects, 13 pairs (43.3%) drew diagrams that represented different types of models from each other; 17 pairs (56.7%) represented institutional repositories using the same type of model. Among those 17 cases, 4 pairs developed *Interface models* and 13 pairs constructed *Processing models*. Subjects developing the *Global-view* models or *Interactivity models* were more likely to disagree with their partners. For instance, subject S3A represented an interactivity model emphasizing human interactions with an institutional repository while subject S3B drew a global-view model. S5A focused on the processes underlying an institutional repository system while his partner subject S5B focused on interactivity between a user and the system.

There were only four cases in which one member of the pair decided to completely redraw the diagram done by his or her partner. These are interesting because they show how subjects can have very different ideas about how an institutional repository works. Strong metaphors sometimes worked against agreement because it was more likely that the partner did not agree with the particular metaphor. For instance, when S8A represented his mental model using a house, S8B disagreed with S8A's interaction model of Ohio State University's Knowledge Bank (see Figures 9-10). Here is the conversation:

Subject S8A: "It's like a house. So here's like the front door, once you get into the door you either will ask a search box and he will tell where to go or you can go this other room and there's a sign 'go here' that will tell you which floor you can go, so you take the little elevator or you can take the stairs. You can either tell the guy, the search box, or you can go to this browse section."

Subject S8B: "I have a totally different representation. I was thinking more like a flow. Making some sort of computer program that search keywords in various documents, all the stuff coded into site, information is coded into site. User can

either search or browse. When you search, system scan all the information and it will give you results. When browsing, user clicks on interesting stuff."

The analysis of conversations between the paired subjects reveals that the partners' comments especially helped to remind their partners of system features that subjects simply missed while drawing. During the explanations and clarifications between the subject pairs, the subjects' recall of and experiences with the institutional repository became richer and additional features were added to the diagrams. The benefits of using this co-discovery method for mental model construction reside in its power to evoke collaborative recall between the paired subjects.

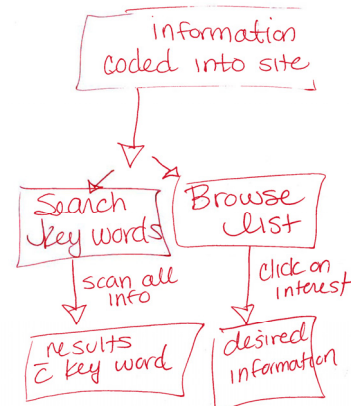
Figure 9. S8A's drawing of Knowledge Bank

Figure 10. S8B's redrawing of S8A's drawing

4. DISCUSSION

The purpose of this study was two-fold: (1) to investigate how undergraduate students construct mental models of new and unfamiliar information systems based on their first experience with the system and (2) to test whether the co-discovery method, in which paired subjects collaborate on tasks in an information retrieval experiment, provides a viable research design to elicit mental models. We have four major findings related to these purposes. First, the results of this study reveal that a majority of subjects depicted a system processing-based mental model to explain how an institutional repository works. The *Interface model* was the next most frequently drawn type of mental model. Few subjects emphasized the *Global-view* or *Interactivity* models. These results differ from Zhang's [1] study which examined undergraduate students' mental models of the Web. Most students in Zhang's study viewed the Web as a "connection between information, people, computers, mobile phones, and webpages" (p. 1334). The next most frequently drawn model in Zhang's study was the "Functional view" in which students saw the Web as a place for doing various activities, such as shopping, entertainment, emailing, paying bills, looking for information, and doing research. Compared to Zhang's 44 students, the subjects in our study expressed more uncertainty about the functionality and features of institutional repositories. As a result, our subjects constructed their mental models focusing more on system operation and interface design rather than on human interactivity with the system or relationships between the institutional repository and other information systems.

We gave our subjects 10 minutes to use each new system. Given this time frame, only 10 out of 60 subjects (16.7%) were able to fully understand the nature and scope of the system. Nine subjects (15.5%) misunderstood aspects of the institutional repository and a quarter of the subjects (25%) gained no understanding of institutional repositories. These finding are consistent with those from an earlier study on experienced institutional repository searchers' perceptions and searching behavior [11]. In St. Jean et al.'s [11] study participants conceived an institutional repository as "database, drawer, receptacle, gateway, interface, place, server, promo, and online forum" or "Wikipedia" (p. 8). In both studies, a majority of users had difficulty in conceptually understanding the institutional repository because it was a new system. When encountering a new and unfamiliar system, subjects across these two studies immediately located the search box and entered a query simply assuming it was "a database of research" [11] rather than exploring the nature, purpose, and scope of the new system to gain some degree of understanding.

In terms of mental model construction, the most common descriptor that subjects used in their drawings was "search" (97 times). The implication of this finding needs to be discussed with caution because of the research design in which subjects were instructed to conduct search tasks before drawing their representations of the system. Still, it is worthwhile noting that "search" was used more frequently than any other descriptors including all the genre/form terms combined (47), metadata-related terms (42) or collection-related terms (18). This finding aligns with the high frequency of the *Processing model* as a means of depicting institutional repositories. Another interesting finding is the low frequency of descriptors that had something to do with institutional units. Thirty-nine out of the total of 366 descriptors were about university, department, community, center, and other academic units. About a half of subjects (n=31)

remembered and wrote down the institutional repository names, such as DSpace, iDEA, Knowledge Bank, Deep Blue, or eScholarship on their drawings. Taking these findings together, we want to highlight the differences between the ways in which users and staff describe an institutional repository. Repository staff members focus more on access to digital materials and developing infrastructure for open access while rarely discussing searching and interaction related elements [9].

The co-discovery method for studying how two people collaborate to find the solution to a common task is referred to by many different names, such as co-participation, paired-user testing, friendly dyads, or constructive interaction [14]. This method has been used in the field of Human Computer Interaction to help "system designers determine whether the basic concepts underlying a system are well understood by users" (p. 27) [14]. We wanted to test whether this method is viable in investigating the construction of mental models in the context of information retrieval experiments. Our study subjects shared one laptop and worked side-by-side in the same room, but only one subject sat in the "driver's" seat using the keyboard and mouse at a time. The partner observed and commented. After one search task was completed, they switched roles. One problem with 'think-aloud' protocols is that the researcher often has to prompt the subject to continue speaking. We expected that through co-discovery we could generate a more natural conversation in each pair. The key factor in this design was that we wanted the paired subjects to explain their experiences, actions, perceptions, and any other ideas to each other, thus eliciting tacit knowledge and providing an audience for the articulation of thoughts throughout the experiment. The results are very encouraging. Paired subjects shared their opinions and ideas during the drawing and searching sessions. While subjects did not revise their original mental models substantially based on their partner's suggestions and comments, they accepted amendments to their drawings made by the partners. Furthermore, the partners often helped each other recall more details from the search tasks that were then incorporated into their mental model drawings.

5. CONCLUSION

The results of this study have implications for two distinct research and practitioner communities. For the interactive information retrieval community, the findings reveal that familiarity with and confidence in using one particular information retrieval system, such as Google, dominates people's approaches to new IR systems and decreases their willingness to learn and explore new systems. Rather than trying to understand the nature and scope of a new system, people simply begin entering keywords into the search box without any prior exploration of or orientation to the system. As a result, mental models of these new systems focus on system processing and interface design, not new types of interactivities. For practitioners and researchers building and studying institutional repositories, the findings of this study suggest that the current emphasis on digital content and open access to introduce and explain institutional repositories needs to be revisited. While the content elements are partially represented, the open access aspect is not readily apparent. Novel ways of helping users to better understand the nature of institutional repositories need to be pursued. One effective way might be to compare similarities and differences between institutional repositories and other types of information retrieval systems with which users are already familiar and better

illustrate the unique aspects of institutional repositories among the multitude of search engines and databases that are available.

7. REFERENCES

[1] Zhang, Y. 2008. Undergraduate students' mental models of the Web as an information retrieval system. *J. Am. Soc. Inf. Sci. Technol.* 59, 13, 2087-2098.

[2] Slone, D. J. 2002. The influence of mental models and goals on search patterns during web interaction. *J. Am. Soc. Inf. Sci. Technol.* 53, 13, 1152-1169.

[3] Norman, D. 1983. Some observations on mental models. In *Mental Models*, D. Gentner and A. Stevens, Eds. Lawrence Erlbaum Associates, Hillsdale, NJ, 7–14.

[4] Fein, R. M., Olson, G. M., and Olson, J. S. 1993. A mental model can help with learning to operate a complex device. In *INTERACT '93 and CHI '93 Conference Companion on Human Factors in Computing Systems* (Amsterdam, The Netherlands, April 24 - 29, 1993). S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel, and T. White, Eds. CHI '93. ACM, New York, NY, 157-158.

[5] Jonassen, D. H. 1995. Operationalizing mental models: strategies for assessing mental models to support meaningful learning and design-supportive learning environments. In *First international Conference on Computer Support for Collaborative Learning* (Indiana Univ., Bloomington, Indiana, United States). J. L. Schnase and E. L. Cunnius, Eds. Computer Support for Collaborative Learning. Lawrence Erlbaum Associates, Hillsdale, NJ, 182-186.

[6] Young, R.M. 1983. Surrogates and mappings: two kinds of conceptual models for interactive devices. In *Mental Models*, D. Gentner and A.L. Stevens, Eds. Lawrence Erlbaum Associates, Hillsdale, NJ, 35-52.

[7] Borgman, C. L. 1985. The user's mental model of an information retrieval system. In *Proceedings of the 8th Annual international ACM SIGIR Conference on Research and Development in information Retrieval* (Montreal, Quebec, Canada, June 05 - 07, 1985). SIGIR '85. ACM, New York, NY, 268-273.

[8] Hendry, D. G. and Efthimiadis, E. N. 2008. Conceptual models for search engines. In *Web Searching: Interdisciplinary Perspectives*. A. Spink & M. Zimmer, Eds. Dordrecht: Springer, 277-307.

[9] Rieh, S. Y., St. Jean, B., Yakel, E., Markey, K., and Kim, J. 2008. Perceptions and experiences of staff in the planning and implementation of institutional repositories. *Library Trends*, 57, 2, 168-190.

[10] Markey, K. Rieh, S. Y., St. Jean, B., Kim, J., and Yakel, E. 2007. *Census of Institutional Repositories in the United States: MIRACLE Project Research Findings*. Washington, D.C.: Council on Library and Information Resources. http://www.clir.org/pubs/reports/pub140/pub140.pdf

[11] St. Jean, B., Rieh, S. Y., Yakel, E., and Markey, K. In press. Unheard voices: Institutional repository end-users. *College & Research Libraries*. http://www.ala.org/ala/mgrps/divs/acrl/publications/crljournal/preprints/crl-071.pdf

[12] Rowlands, I., Nicholas, D., Williams, P., Huntington, P., Fieldhouse, M., Gunter, B., et al. 2008. The Google generation: the information behaviour of the researcher of the future. *Aslib Proceedings*, 60, 4, 290-310.

[13] Urquhart, C. and Rowley, J. 2007. Understanding student information behaviour in relation to electronic information services: lessons from longitudinal monitoring and evaluation. *J. Am. Soc. Inf. Sci. Technol*. 58, 8, 1188-1197.

[14] Kahler, H. 2000. Constructive interaction and collaborative work: introducing a method for testing collaborative systems. *Interactions* 7, 3 (Jun. 2000), 27-34.

[15] O'Malley, C., Draper, S., & Riley, M. 1984. Constructive interaction: A method for studying human-computer-human interaction. In *Proceedings of INTERACT'84: 1st IFIP International Conference on Human-Computer Interaction*. Amsterdam: Elsevier, 1-5.

[16] Proffitt, M. 2007. How and why of user studies: RLG's RedLightGreen as a case study, *Journal of Archival Organization*, 4, 1, 87-110.

[17] Borlund, P. 2003. The IIR evaluation model: A framework for evaluation of interactive information retrieval systems. *Inform. Res.* 8, 3. http://informationr.net/ir/8-3/paper152.html

[18] Schonfeld, R. C. and Housewright, R. 2010. Faculty survey 2009: key strategic insights for libraries, publishers, and societies. http://www.ithaka.org/ithaka-s-r/research/faculty-surveys-2000-2009/Faculty%20Survey%202009%20CH1.pdf