Interdisciplinary Curriculum Development for Digital Library Education

Seungwon Yang, Edward A. Fox,
Department of Computer Science, Virginia Tech,

Barbara M. Wildemuth, Jeffrey Pomerantz,
and Sanghee Oh
School of Information and Library Science, University of North Carolina – Chapel Hill

Acknowledgements (Selected)

- Sponsors
  - National Science Foundation

- Faculty / Staff
  - Edward A. Fox, Barbara M. Wildemuth, Jeffrey P. Pomerantz, Sanghee Oh, John Moore, Melanie Darden, …

- Students
  - Uma Murthy, Seonho Kim, Xiaoyan Yu, Ryan Richardson, …

Overview

- Introduction
- Foundations
- Module development
- Invitation

Introduction

- Project description
- Problems, challenges, motivation
- What we do
- How we do it

Project Description

- Joint project: VT CS & UNC-CH SILS

- NSF Grant:

  VT(IIS-0535057) & UNC-CH(IIS-0535060)

- Period

  Jan. 2006 - Dec. 2008

Problems, Challenges, Motivation

- Organizing DL topics into groups
  - → Different groupings possible
  - → Granularity of groups
- Resource preparation
  - → Collect, develop, use publicly available ones
  - → Copyright issues

What we do

- Develop DL lessons (modules)
  - 10 core and their sub-modules
  - Fox (PI) & Gonçalves textbook (in progress)
- NSDL service
  - Modules will be available
  - Instructors create courses from them

How we do it

- Define → develop → test (2 cycles)

- Guidances
  1. Advisory board
  2. Doctoral consortium at JCDL sponsored by IEEE TCDL
  3. 5S framework
  4. Computing Curriculum 2001

Overview

- Introduction
- Foundations
- Module development
- Invitation

Foundations

- Advisory board
- 5S framework
- Computing Curriculum 2001

Advisory Board

- DL researchers & practitioners

- Meetings:
  - 5/1/06 at VT
  - 6/13/06 at UNC-CH (during JCDL '06)
  - 11/6/06 at ASIST '06

Advisory Board Members

From VT	From UNC	Other Institutions
Steven Edwards, CS	Catherine Blake, SILS	Dan Atkins, U. Michigan
Roger Ehrich, CS	Laura Gasaway, Law School	Christine Borgman, UCLA
Weiguo Fan, ACIS	Jane Greenberg, SILS	Lillian Cassel, Villanova
Steve Harrison, CS	Stephanie Haas, SILS	Michael Christel, CMU
Gail McMillan, Library	Brad Hemminger, SILS	Raya Fidel, U. Washington
Chris North, CS	Thomas James, Dean, School of Education	Richard Furuta, Texas A&M University
Manuel Pérez-Quiñónez, CS	Paul Jones, Director, ibiblio; SILS & School of Journalism & Mass Communication	Elizabeth Liddy, Syracuse University
Naren Ramakrishnan, CS	Diane Kelly, SILS	Clifford Lynch, CNI
Deborah Tatar, CS	Gary Marchionini, SILS	Kurt Maly, ODU
Layne Watson, CS	Montek Singh, CS	Javed Mostafa, Indiana
	Natasha Smith, Library	Tefko Saracevic, Rutgers
	Helen Tibbo, SILS	Linda Smith, UIUC
	Steve Weiss, CS	Ingeborg T. Sølvberg, NTNU (Norway)

Selected International Collaborators

- India

  Arun Kumar Chakraborty

  (Bose Institute, Kolkata, India)

- Brazil

  Ricardo da Silva Torres

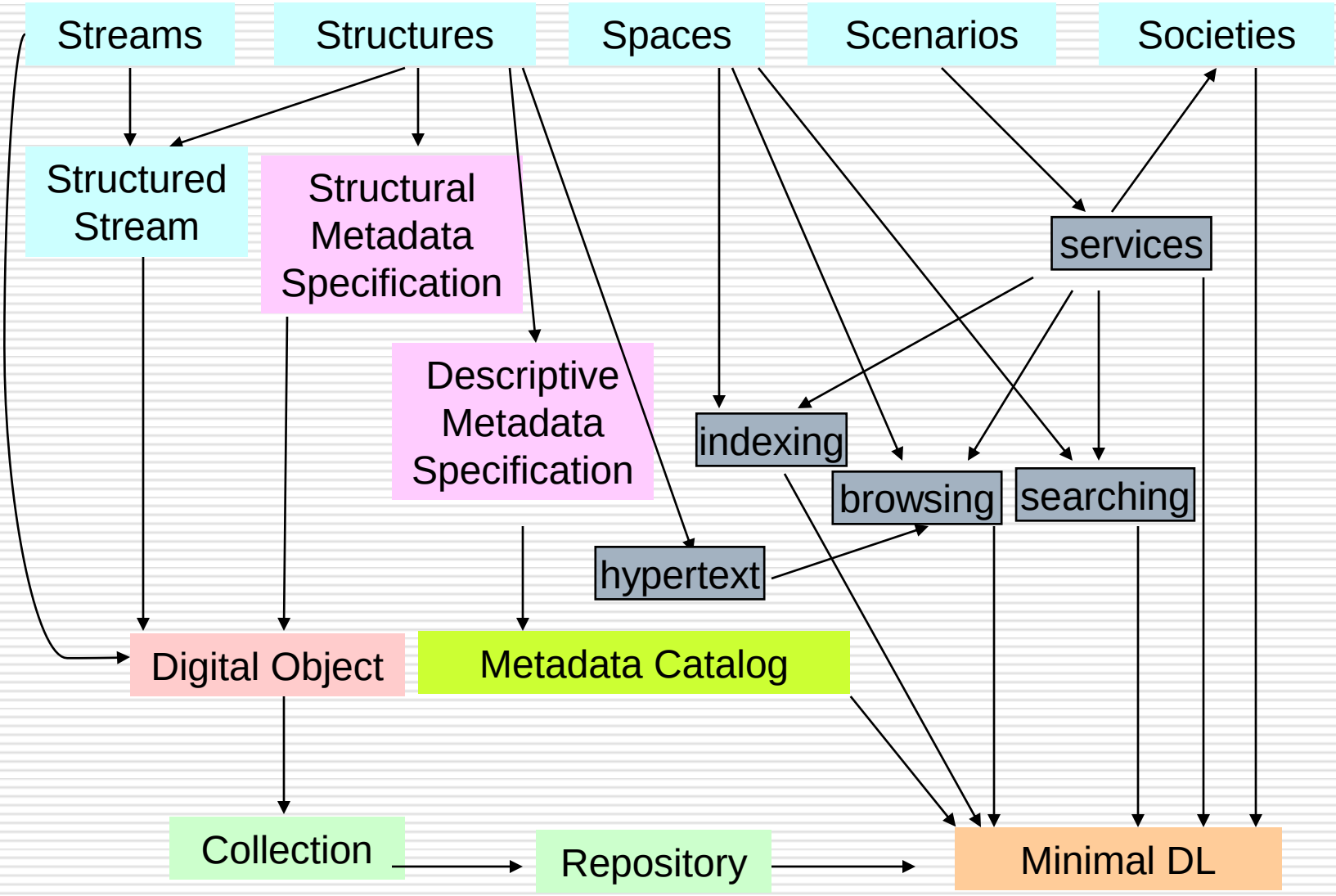
  (State University of Campinas (UNICAMP), SP, Brazil)

5S Framework

- Theoretical foundation
- Necessary condition to be a Minimal DL

S's	**Role in DL**	**Examples**
Streams	Various types of content	text, video, audio
Structures	Organize information	catalog, metadata, hypertext
Spaces	Store and present information	interfaces, storage, vector space
Scenarios	Provide services	searching, browsing, recommending
Societies	Form a DL community of users and service managers	service managers, teachers, learners

A Minimal DL in the 5S Framework



Shen, R. et al.

Computing Curriculum 2001

- CS body of knowledge
  → Information Management (IM.14)
  → Digital Libraries

- DL topics and learning objectives explained

Overview

- Introduction
- Foundations
- Module development
- Conclusion

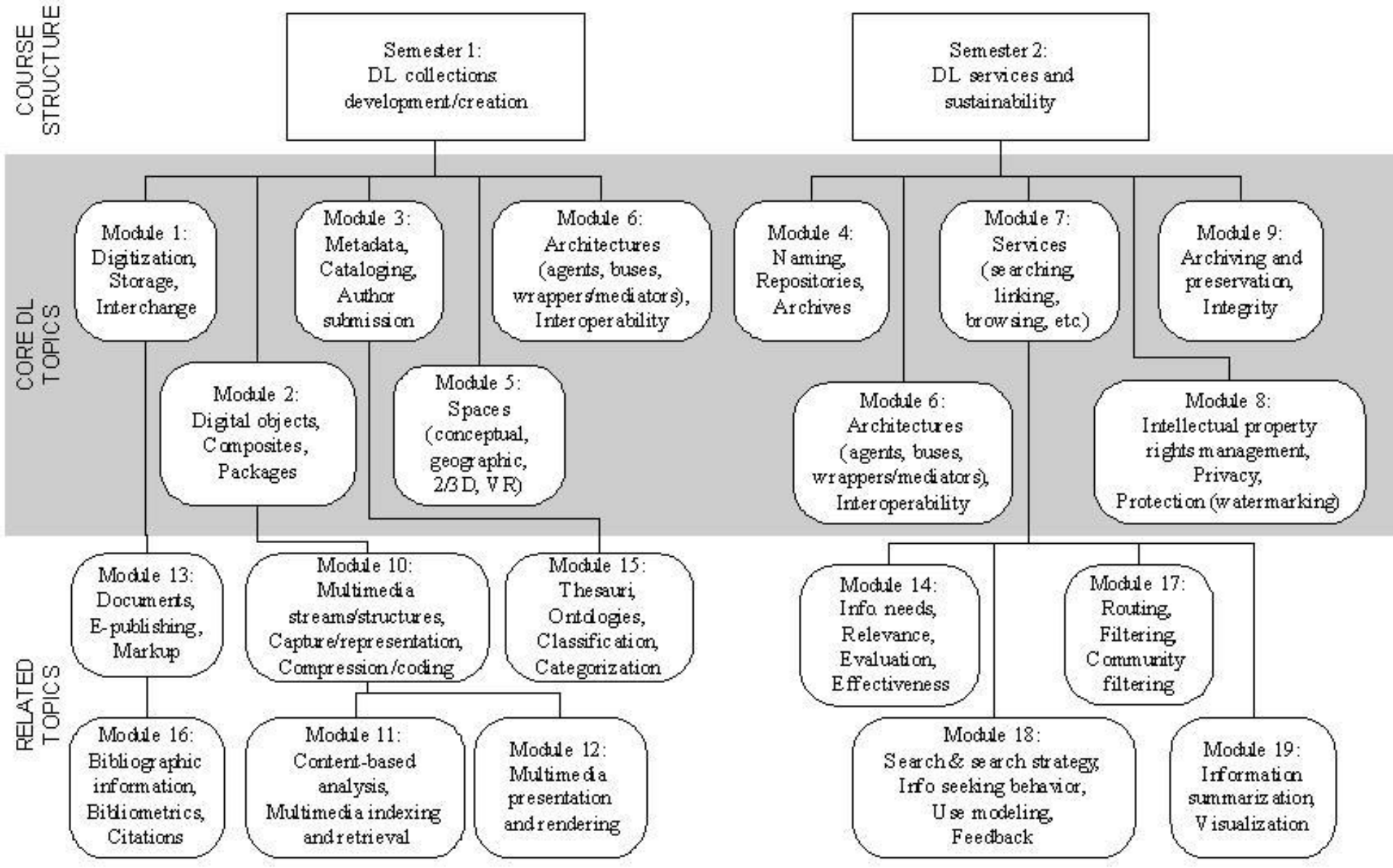
Module Development

- DL module set
- Why modular design
- Module template
- Draft module examples

Module Set

- Initial 19 → 10 core & sub-modules
- Some integrated into others
- New ones added

DL Topics in 19 Modules (original)



DL Literature Analyses

- ACM DL, JCDL, and D-Lib Magazine articles (approx. 1100)
  - Published in JCDL ’06 proceedings
  - Most frequent topics
  - Used old version of module set

Modules (old version)

1. Collection Development
2. Digital objects/Composites/Packages
3. Metadata, Cataloging, Author submission
4. Architecture, Interoperability
5. Data visualization
6. Services
7. Intellectual property rights management, Privacy, Protection
8. Social issues / Future of DLs
9. Archiving and Preservation

Distribution of Conference Paper Topics

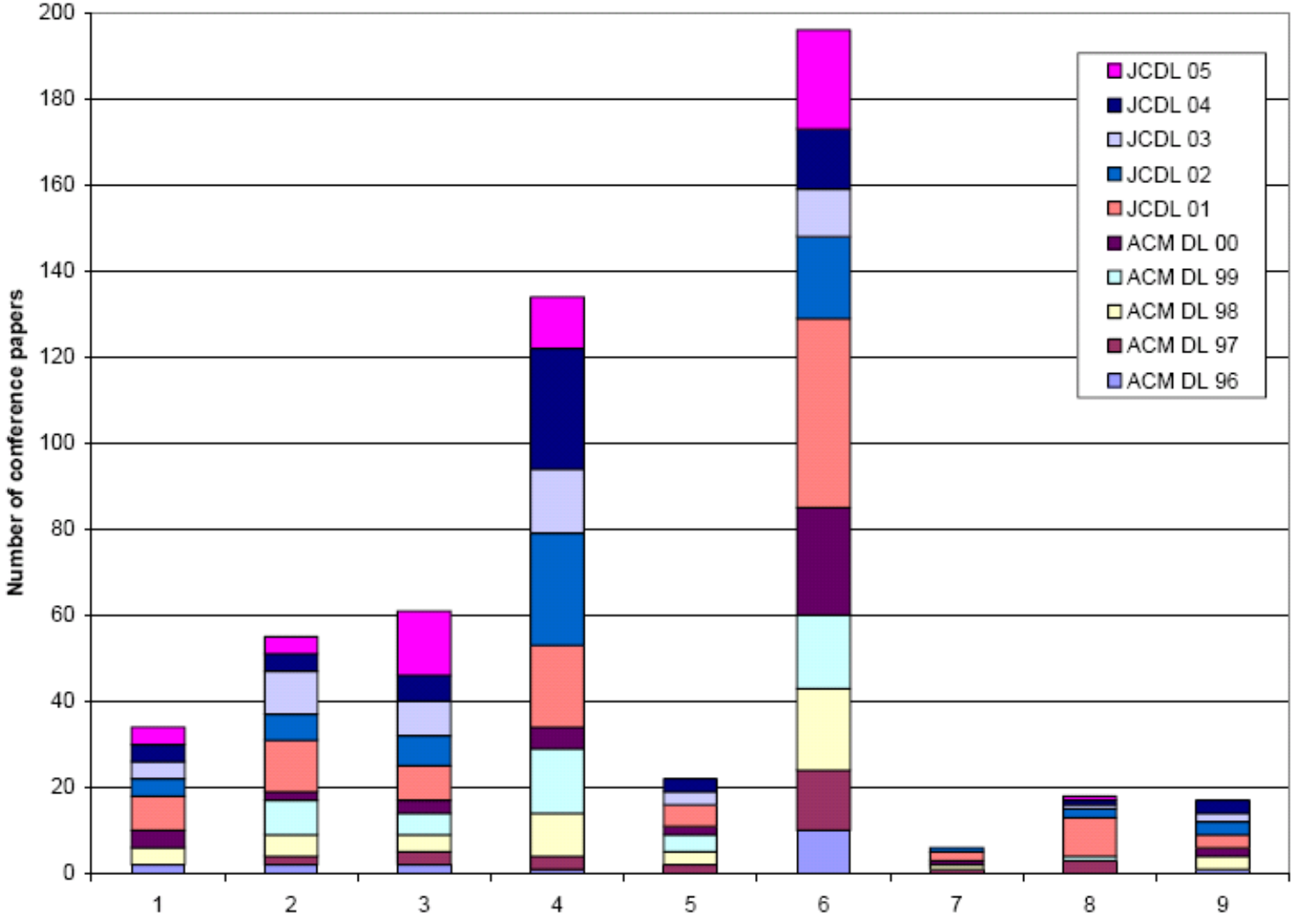


Figure 2. Distribution of conference papers across module topics

Distribution of D-Lib Magazine Paper Topics

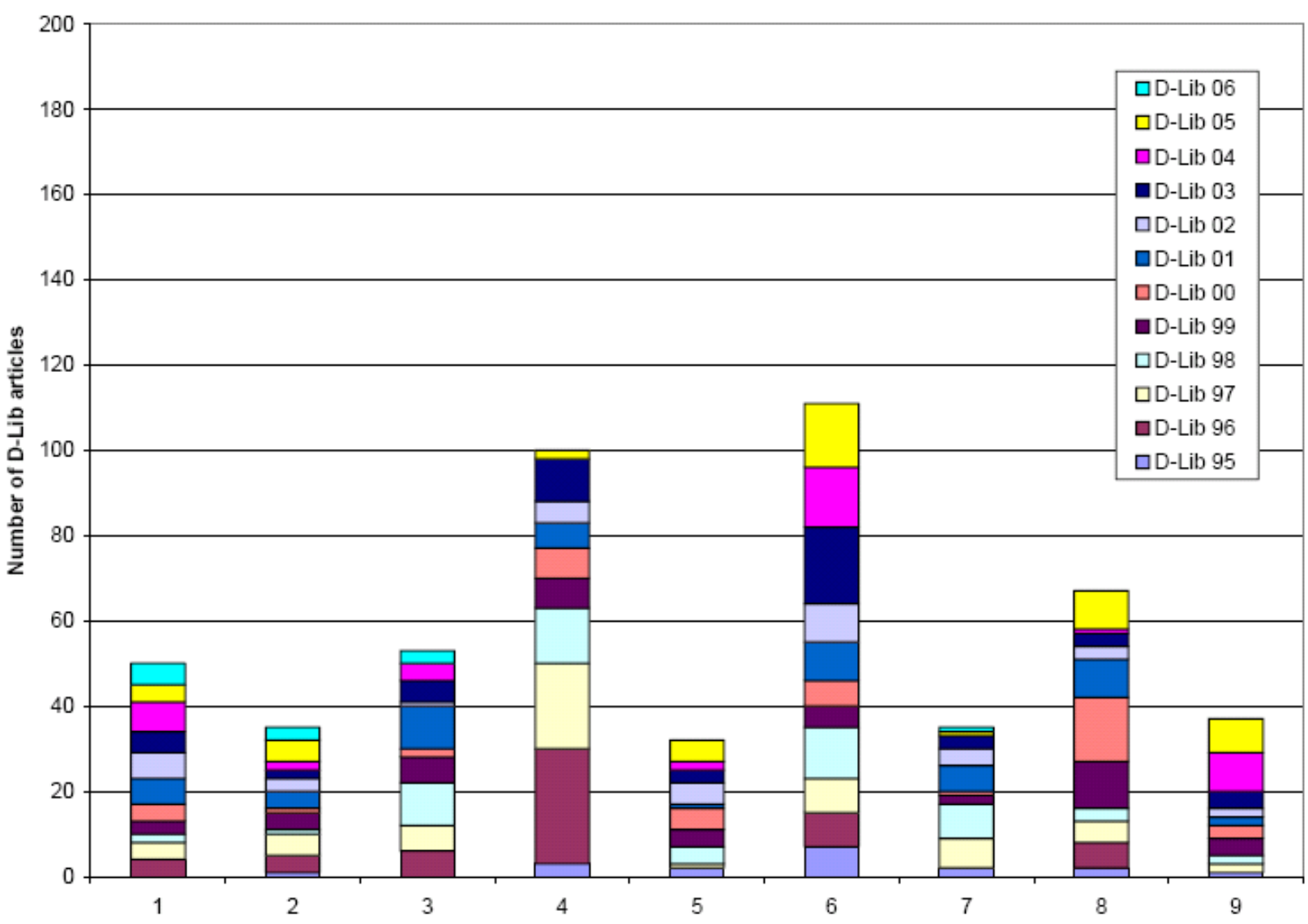
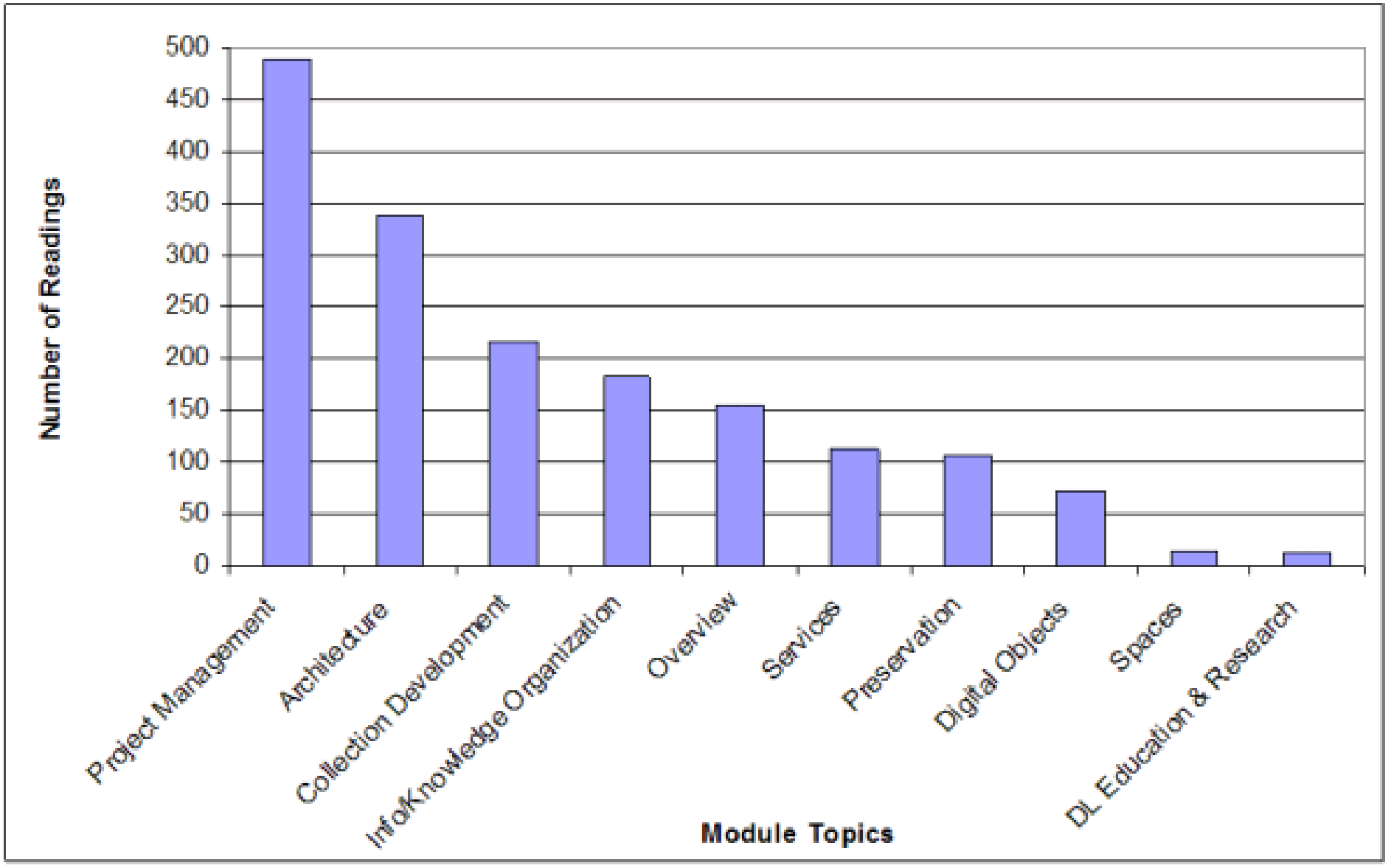


Figure 3. Distribution of D-Lib Magazine papers across module topics

DL Course Syllabi Analyses

- All readings in DL courses in LIS programs accredited by the American Library Association (approx. 1600)
  - Published in D-*Lib Magazine* (Nov. 2006)
  - Most frequent topics
  - Most frequently-assigned authors, journal articles, books, and journal titles

DL Syllabi Analyses (LIS area)
- Distribution of Reading Topics -



Most Frequently Assigned Books (selected)

Books	# of assignments
Witten, I. H., & Bainbridge, D. (2003). How to Build a Digital Library. San Francisco, CA: Morgan Kaufman Publishers.	10
Arms, W. Y. (2000). Digital Libraries. Cambridge, MA: The MIT Press.	9
Borgman, C. L. (2000). From Gutenberg to the Global Information Infrastructure. Cambridge, MA: The MIT Press.	7
Lesk, M. (2004). Understanding Digital Libraries (Second ed.). San Francisco, CA: Morgan Kaufman Publishers.	6
Chowdhury, G. G., Chowdhury, S. (2003). Introduction to Digital Libraries. London: Facet.	5

Most Frequently Assigned Journal Articles (selected)

Articles	# of assignments
Borgman, C. L. (1999). What are Digital Libraries? Competing Visions. Information Processing & Management, 35(3), 227-243.	8
Bush, V. (1945). As We May Think. The Atlantic Monthly, 101-108.	6
Arms, W. Y., Blanchi, C., & Overly, E. A. (1997). An Architecture for Information in Digital Libraries. D-Lib Magazine, 3(2).	5
Schwartz, C. (2000). Digital Libraries: An Overview. Journal of Academic Librarianship, 26(6), 385-394.	5
McCray, A. T., Gallagher, M. E. (2001). Principles for digital library development. Communications of the ACM, 44(5), 48-54.	5
Lossau, N. (2004). Search Engine Technology and Digital Libraries: Libraries Need to Discover the Academic Internet. D-Lib Magazine, 10(6).	5
Lynch, C. (2005). Where Do We Go From Here? The Next Decade for Digital Libraries. D-Lib Magazine, 11(7/8).	5
Hill, L. L., Carver, L., Larsgaard, M., Dolin, R., Smith, T. R., Frew, J. (2000). Alexandria digital library: User evaluation studies and system design. Journal of the American Society for Information Science, 51(3), 246-259.	4

Most Frequently Assigned Journals (selected)

Journal names	# of assignments
D-Lib Magazine	127
Communications of the ACM	27
First Monday	23
Journal of the American Society for Information Science (& Technology)	22
Computers in Libraries	16
Information Processing & Management	14
Ariadne	13
Library Trends	13
Online	12
RLG DigiNews	10

Most Frequently Assigned Authors (selected)

Author	# of assignments	# of unique works assigned
Arms, William Y.	30	11
Jasco, Peter	28	28
Borgman, Christine L.	24	10
Witten, Ian H.	21	9
Bainbridge, David	20	8
Lynch, Clifford	19	11
Kenney, Anne R.	18	9
Lagoze, Carl	17	10
Marchionini, Gary	15	12
Smith, Abby	14	6

Ongoing Discussion of Concepts/Words from JCDL, ACM DL, and D-lib Articles

- Data conversion vs. digitization
- Acquisition vs. harvesting
- Text summarization vs. information summarization
- Document disaggregation vs. transformation

- information management - under ‘services’
- online publication (e-publishing) - under ‘services’
- policy - controversial
- personalization - under ‘services’
- Hypertext, hypermedia – under ?

Missing Concepts/Words found in JCDL, ACM DL, and D-lib Articles

- We added a module to cover concepts related to the surrounding environments of DL
  - DL applications (e.g., DL use in education)
  - DL design/user studies/evaluation/appliances /hardware (e.g., XLibris - active reading machine)
  - DL management
  - E-commerce, pricing of DL contents
  - Training D-librarians

SEMESTER 1: DL COLLECTIONS: DEVELOPMENT, CREATION

CORE TOPICS

1	**Overview**
2	**Collection Development**
3	**Digital Objects**
4	**Info/ Knowledge Organization**
5	**Architecture (agents, mediators)**

2-a: Digitization
2-b: Document and e-publishing markup
2-c: Harvesting

3-a: Text resources
3-b: Multimedia
3-c (8-a): File formats, transformation

4-a: Metadata, harvesting, cataloging
4-b: Ontologies, classification, categorization
4-c: Vocabulary control (e.g., thesauri, terminologies, etc.)
4-d: Bibliographies, bibliometrics, Webometrics
4-e: Indexing

5-a: Interoperability
5-b: Sustainability
5-e: Identifiers, handles, DOI, PURL
5-h: Applications (e.g., Greenstone, Fedora, Dspace)
5-i: Web publishing (e.g., wiki, rss, Moodle, etc.)

SEMESTER 2: DL SERVICES

CORE TOPICS

5	**Architecture (agents, mediators)**
6	**Spaces (conceptual, geo-graphic, 2/3D, VR)**
7	**Services**
8	**Archiving and preservation integrity**
9	**Project management**
10	**DL education and research**

5-c: Interface design, usability assessment
5-d: Search engines, IR
5-f: Info summarization, visualization
5-g: Recommender systems
5-j: Security

6-a: Storage
6-b: Repositories, archives

7-a: Info needs, relevance, evaluation
7-b: Search strategy, info seeking behavior, user modeling
7-c: Reference services
7-d: Routing, community filtering
7-e: Sharing, networking, interchange (e.g., social)

8-a, 3-c: File formats, transformation

9-a: DL development for a specific domain
9-b: DL project examples
9-c: DL evaluation
9-d: Legal issues (e.g., copyright)
9-e: Cost/economic issues
9-f: Social issues

10-a: Future DLs

Why Modular Design

- Flexibility
- E.g.,
  - A course based on a single module
  - Course sequence (program) from multiple modules
  - Plug in modules into an existing course (enhancement)
    - Module 1. Overview / Module 10. DL Education & Research

Module Template Design

- Based on…
  - Educational experience of the research team
  - CC2001 (title, hours, topics, learning objectives)
  - Advisory board meeting discussions

Module Template (1/2)

1. **Module name**
2. **Learning objectives**
3. **Level of effort required**
   (in and out-of-class time of students)
4. **Prerequisite knowledge required**
   (completion optional)
5. **Relationships with other modules**
   (flow between modules)
6. **5S characteristics of the module**
7. **Introductory remedial instruction**
   (completion optional; intended to address the prerequisite knowledge/skills required)
8. **Resources**
   (all the resources in the 'Body of knowledge' section)

Module Template (2/2)

9. Body of knowledge (Theory + Practice)

Topics might be skipped or studied in different orders

Topic 1

Theories and background knowledge of the topic

Learning activities

Presentation slides

Interactive demo

Resources

- Textbooks (one or multiple chapters might be assigned)
- Reference papers (relevant parts might be marked with SI tool)
- Advanced reading

Worksheets

Topic 2 ...

Topic 3 ...

10. Concept maps (created by students)

11. Exercises / Learning activities

12. Evaluation of learning outcomes

13. Glossary

14. Useful links

Draft Module Example
Module 3_b:Multimedia

5. Digital Audio Processing (20pts)

-Pretest (grade does not count)

-Study online text **Chapter.5** but you can skip 5.2.4 and 5.2.5 and 5.2.6

-Carry out activities

Personal and group concept maps (1 pt)

Interactive tutorial on audio dithering

Worksheet on audio dithering, where you should turn in the Matlab log, but only do Exercises 1, 2, 3, 4 and Questions 1 and 2. **worksheet.1** (3 pts)

Worksheet on digital audio file size and file transfer time: **worksheet.2** (3 pts)

Interactive tutorial on non-linear quantization and mu-law encoding

Worksheet on non-linear companding and mu-law encoding: **worksheet.3** (3 pts)

Please follow the instructions in **Submission Instructions** to submit the worksheets

-Posttest (grade counts to demonstrate level of mastery) (10 pts)

Interactive Tutorial Example

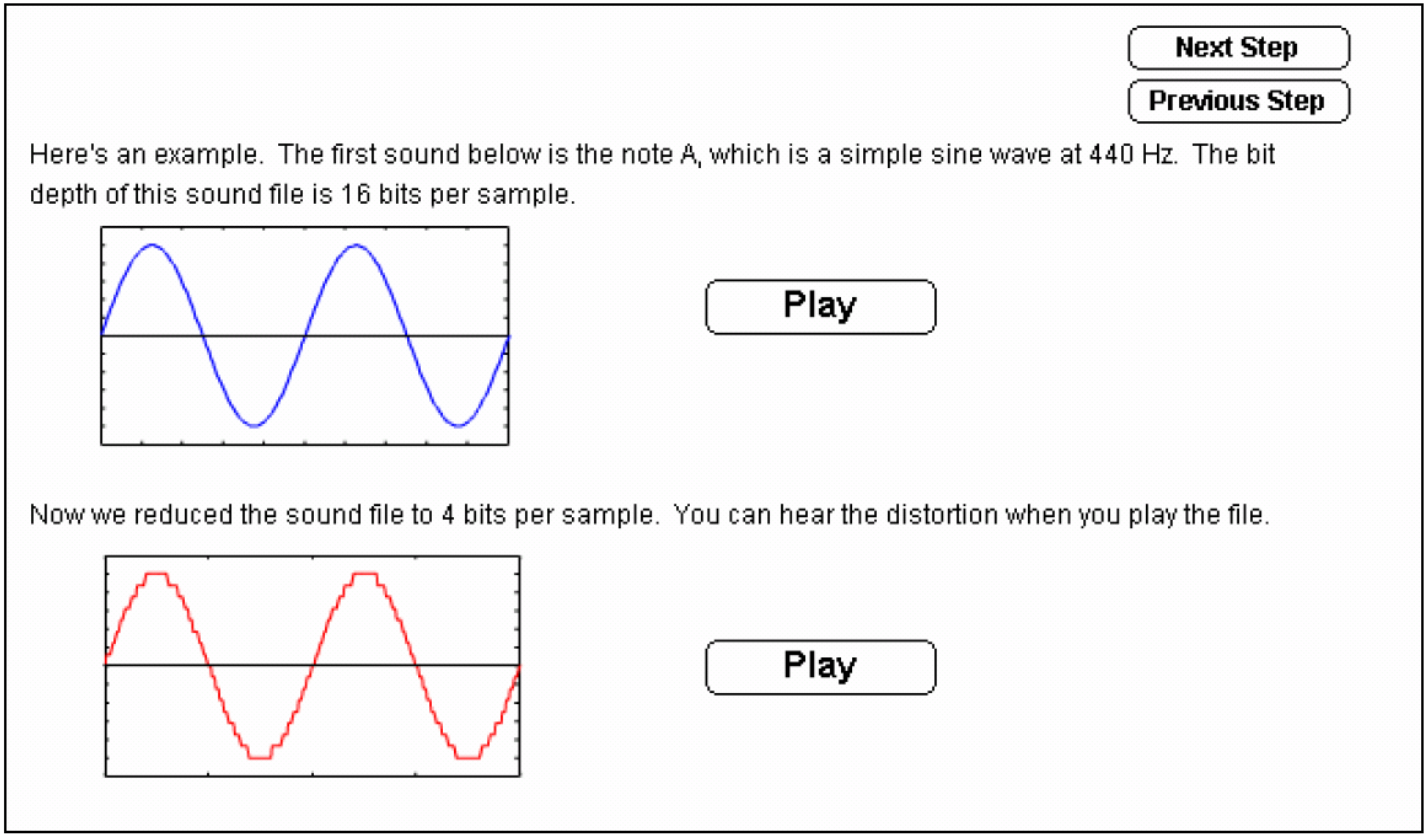


Fig. 3. An example of interactive tutorial, 'Audio Dithering'

Draft Module Example

Module 7-a:Info. needs, relevance, evaluation/effectiveness

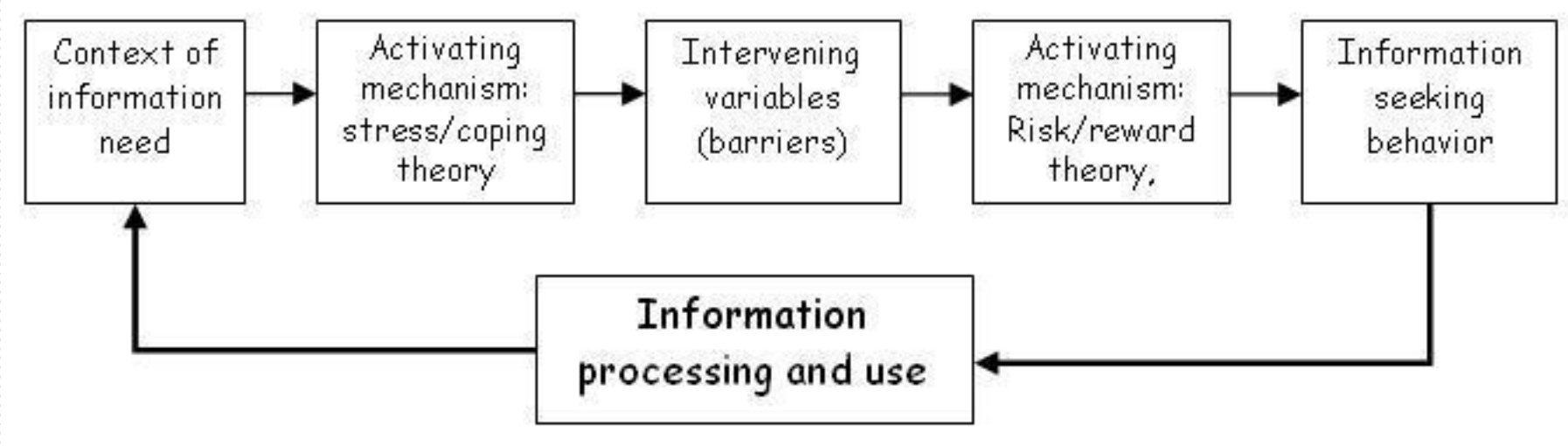
Wilson, P. (1973). Situational relevance. Information Storage & Retrieval, 9, 457-471.

Wilson, T. D. (1997). Information behaviour: An interdisciplinary perspective. Information Processing & Management, 33(4), 551-572.

Yuan, X.-J., Belkin, N. J., & Kim, J.-Y. (2002). The relationship between ASK and relevance criteria. Proceedings of SIGIR 2002, 359-360.

c. Outline of knowledge to be covered

WILSON'S MODEL OF INFORMATION BEHAVIORS



Learning Activity Example

Discussion activity: Personal experiences of an information need

To immediately follow the review of Wilson's generalized model of information behavior

Students in the class should be formed into pairs. In each pair, one student will interview the other. (This process should later be repeated, reversing roles.) The person being interviewed should be asked to recall a recent experience of having an information need. The interviewer should ask about the content of the information need, the context in which it arose, and the process through which it was pursued (successfully or unsuccessfully). The pair should then evaluate what was learned about this example of an information need and see if Wilson's model fully describes the process. Were there aspects of the information-seeking episode that are not covered in Wilson's model? Are there aspects of Wilson's model that did not occur during this information-seeking episode?

Fig. 5. A sample learning activity from module 14

In Each Module Content…

- Clear mapping required
  - Each course objective → corresponding lecture, resources, and exercises
  - Organized module
  - Helps students' understanding

Overview

- Introduction
- Foundations for curriculum development
- Digital library module development
- Invitation

Invitation

- Work together, get benefits together
- Join – two ways
  - Educational module development
  - Course/modules evaluation
- Contact information
  - Edward A. Fox fox@vt.edu

Resources

- Project homepage
  http://curric.dlib.vt.edu

- Wiki
  http://curric.dlib.vt.edu/wiki
  publications, advisory board, modules, project progress diary, etc.

Thank you!
Questions?

CC 2001 Information Management(IM)

IM14. Digital libraries [elective] *Topics:*

1. **Digitization, storage, and interchange**
2. **Digital objects, composites, and packages**
3. **Metadata, cataloging, author submission**
4. **Naming, repositories, archives**
5. **Spaces (conceptual, geographical, 2/3D, VR)**
6. **Architectures (agents, buses, wrappers/mediators), interoperability**
7. **Services (searching, linking, browsing, and so forth)**
8. **Intellectual property rights management, privacy, protection (watermarking)**
9. **Archiving and preservation, integrity**

Learning objectives:

1. **Explain the underlying technical concepts in building a digital library.**
2. **Describe the basic service requirements for searching, linking, and browsing.**
3. **Critique scenarios involving appropriate and inappropriate use of a digital library, and determine the social, legal, and economic consequences for each scenario.**
4. **Describe some of the technical solutions to the problems related to archiving and preserving information in a digital library.**
5. **Design and implement a small digital library.**

Textbook on DLs

- Rely on the 5S framework
  - Integrated coverage of the many concepts related to DLs
- Book for teaching & reference

Textbook Outline (draft)

- Ch. 1. Introduction (Motivation, Synopsis)
- Part 1 – The “Ss”
  - Ch. 2: Streams
  - Ch. 3: Structures
  - Ch. 4: Spaces
  - Ch. 5: Scenarios
  - Ch. 6: Societies

Textbook Outline (draft)

- Part 2 – Higher DL Constructs
  - Ch. 7: Collections
  - Ch. 8: Catalogs
  - Ch. 9: Repositories and Archives
  - Ch. 10: Services
  - Ch. 11: Systems
  - Ch. 12: Case Studies

Textbook Outline (draft)

- Part 3 – Advanced Topics
  - Ch. 13: Quality
  - Ch. 14: Integration
  - Ch. 15: How to build a digital library
  - Ch. 16: Research Challenges, Future Perspectives
- Appendix
  - A: Mathematical preliminaries
  - B: Formal Definitions: Ss
  - C: Formal Definitions: DL terms, Minimal DL
  - D: Formal Definitions: Archeological DL
  - E: Glossary of terms, mappings

SEMESTER 1: DL COLLECTIONS: DEVELOPMENT, CREATION

CORE TOPICS

1	**Overview**	
2	**Collection Development**	2-a: Digitization 2-b: Document and e-publishing markup 2-c: Harvesting
3	**Digital Objects**	3-a: Text resources 3-b: Multimedia 3-c (8-a): File formats, transformation
4	**Info/ Knowledge Organization**	4-a: Metadata, harvesting, cataloging 4-b: Ontologies, classification, categorization 4-c: Vocabulary control (e.g., thesauri, terminologies, etc.) 4-d: Bibliographies, bibliometrics, Webometrics 4-e: Indexing
5	**Architecture (agents, mediators)**	5-a: Interoperability 5-b: Sustainability 5-e: Identifiers, handles, DOI, PURL 5-h: Applications (e.g., Greenstone, Fedora, Dspace) 5-i: Web publishing (e.g., wiki, rss, Moodle, etc.)

SEMESTER 2: DL SERVICES
CORE TOPICS

5	**Architecture (agents, mediators)**	5-c: Interface design, usability assessment 5-d: Search engines, IR 5-f: Info summarization, visualization 5-g: Recommender systems 5-j: Security
6	**Spaces (conceptual, geo-graphic, 2/3D, VR)**	6-a: Storage 6-b: Repositories, archives
7	**Services**	7-a: Info needs, relevance, evaluation 7-b: Search strategy, info seeking behavior, user modeling 7-c: Reference services 7-d: Routing, community filtering 7-e: Sharing, networking, interchange (e.g., social)
8	**Archiving and preservation integrity**	8-a, 3-c: File formats, transformation
9	**Project management**	9-a: DL development for a specific domain 9-b: DL project examples 9-c: DL evaluation 9-d: Legal issues (e.g., copyright) 9-e: Cost/economic issues 9-f: Social issues
10	**DL education and research**	10-a: Future DLs

(selected) Most Frequently Assigned...

		# of readings assigned
Authors	Arms, William Y.	30
	Jasco, Peter	28
	Borgman, Christine L.	24
	Witten, Ian H.	21
	Bainbridge, David	20
Articles	Borgman, What are digital libraries? Competing visions	8
	Bush, As We May Think	6
	Arms, An Architecture for Information in Digital Libraries	5
	Schwartz, Digital libraries: an overview	5
	Chowdhury, Introduction to digital libraries	5
Books	Witten, How to Build a Digital Library	10
	Arms, Digital Libraries	9
	Borgman, From Gutenberg to the Global Information Infrastructure	7
	Lesk, Understanding Digital Libraries	6
	Lesk, Practical digital libraries: books, bytes, and bucks	4

Table 4. Evaluation questions and data collection activities					
Evaluation questions	**Data collection activities**				
	Content analysis	Open-ended ques-tion-naires	Indi-vidual inter-views	In-class student survey	Course assign-ments, tests, etc.
Pre-implementation inspection by experts					
Does the module adequately cover the topic?	X	X	X		
Are the assigned/suggested readings current and appropriate?	X	X	X		
Does the module address desired skills at the appropriate level of Bloom's taxonomy?	X	X	X		
Post-implementation instructor perceptions					
Did the module adequately cover the topic?		X	X		
Were the assigned/suggested readings current and appropriate?		X	X		
Did students display desired skills at the appropriate level of Bloom's taxonomy?		X	X		
Post-implementation student perceptions					
Were the module content and the readings interesting, useful, and challenging? Was the module structured appropriately?				X	
How much does the student feel he/she has learned from the module?				X	
What is the student's GPA? Anticipated course grade?				X	
Post-implementation student performance					
How did students perform on assessments of their learning of the module's content?					X