



**DIGITAL LIBRARIES**  
Curriculum Development

# Improving Education and Understanding of NDLT

- ETD Symposium 2007 -

Seungwon Yang  
Sanghee Oh  
Jeffrey P. Pomerantz  
Barbara M. Wildemuth  
Edward A. Fox

10th International Symposium on Electronic Theses and Dissertations  
Uppsala, Sweden, 13<sup>th</sup> – 16<sup>th</sup>, June 2007



# Acknowledgement

- This material is based upon work supported by the National Science Foundation under Grant Nos. IIS-0535057 (VT) and IIS-0535060 (UNC-CH)
- Special thanks to
  - ETD 2007 Organizers
  - DL Curriculum Project Advisory Board

# Project Web Site

- [curric.dlib.vt.edu](http://curric.dlib.vt.edu)

- It contains:

- Links to module frameworks, template
    - Links to published papers, articles
    - Project progress diary
    - Advisory board members
    - News and interviews

# Motivation for Scholars

- One of the NDLTD goals is, “for graduate students to learn about **electronic publishing** and **digital libraries**, applying that knowledge as they engage in their **research** and build and submit their own **ETD**” – excerpt from NDLTD site –
- Our project team selected 11 DL educational modules (out of 43 total), which we think are the most relevant to scholars’ research efforts
- By studying those modules, their corresponding ETD guide sections and related resources, scholars might be able to achieve the goal in the excerpt above

# Outline

1. DL educational module framework
2. Selected modules for scholars
3. Selected modules & mapping to ETD guide
4. Selected modules & their resources
5. Summary
6. Future work

# DL Module Framework (1 of 2)

- Core modules -

- 10 core modules (43 sub modules)
  - Original 19 modules → 9 modules  
→ 10 modules → updated 10 modules
  - Module updates
    - A series of papers, articles & syllabi readings classifications
    - New concepts, keywords were discussed and added
    - Some topics were merged together

# DL Module Framework (2 of 2)

- Sub module assignment -

- Sub-modules assigned to more than two cores
  - 1-a(10-c): Conceptual framework, theories, definitions
    - ➔ in 1: Overview and 10: DL education & research
  - 2-c(8-c): File formats, transformation, migration
    - ➔ in 2: Digital objects and 8: Preservation
- Except for the above two sub-modules, each was assigned to only one core

# Core Modules (1 of 2)

## CORE TOPICS

1	Overview	1-a (10-c): Conceptual frameworks, theories, definitions	1-b: History of digital libraries and library automation
2	Digital Objects	2-a: Text resources 2-b: Multimedia	2-c (8-c): File formats, transformation, migration
3	Collection Development	3-a: Collection development/selection policies 3-b: Digitization	3-c: Harvesting 3-d: Document and e-publishing/presentation markup
4	Info/ Knowledge Organization	4-a: Information architecture (e.g., hypertext, hypermedia) 4-b: Metadata, cataloging, metadata markup, metadata harvesting 4-c: Ontologies, classification, categorization	4-d: Subject description, vocabulary control, thesauri, terminologies 4-e: Object description and organization for a specific domain
5	Architecture (agents, mediators)	5-a: Architecture overviews/models 5-b: Applications 5-c: Identifiers, handles, DOI, PURL	5-d: Protocols 5-e: Interoperability 5-f: Security



# Core Modules (2 of 2)

## CORE TOPICS

6	User Behavior/ Interactions
---	--------------------------------

6-a: Info needs, relevance  
6-b: Search strategy, info seeking behavior, user modeling

6-c: Sharing, networking, interchange (e.g., social)  
6-d: Interaction design, info summarization and visualization, usability assessment

7	Services
---	----------

7-a: Search engines, IR, indexing methods  
7-b: Reference services  
7-c: Recommender systems

7-d: Routing, community filtering  
7-e: Web publishing (e.g., wiki, rss, Moodle, etc.)

8	Preservation
---	--------------

8-a: Approaches to archiving and repository development

8-b: Sustainability  
8-c (2-c): File formats, transformation, migration

9	Management and Evaluation
---	------------------------------

9-a: Project management  
9-b: DL case studies  
9-c: DL evaluation, user studies  
9-d: Bibliometrics, Webometrics

9-e: Legal issues (e.g., copyright)  
9-f: Cost/economic issues  
9-g: Social issues

10	DL education and research
----	------------------------------

10-a: Future of DLs  
10-b: Education for digital librarians

10-c (1-a): Conceptual framework, theories, definitions  
10-d: DL research initiatives



# Outline

1. DL educational module framework
2. **Selected modules for scholars**
3. Selected modules & mapping to ETD guide
4. Selected modules & their resources
5. Summary
6. Future work

# Selected Modules for Scholars (1 of 7)

- 11 sub-modules were selected
  - Most relevant to scholars' research endeavor
  - e.g., finding ETDs, preparing ETDs and submitting ETDs to ND LTD, developing DLs
- Complete table is at [curric.dlib.vt.edu](http://curric.dlib.vt.edu)

# Selected Modules for Scholars (2 of 7)

- 1-a(10-c): Conceptual frameworks, theories, definitions
  - Provides the foundational DL concepts to scholars
- 2-c(8-c): File formats, transformation, migration
  - Helps transformation of doc, html and other file formats into PDF for submission to ETD systems (e.g., NDLTD)

# Selected Modules for Scholars (3 of 7)

## ■ 3-b: Digitization

- Small scale – scanning of a picture for an ETD
- Large scale – automatic and systematic digitization process to develop DLs from traditional materials (e.g., books, historical documents, etc.)

## ■ 3-d: Document and e-publishing/presentation markup

- Explains the major components, functions, applications of XML
- Helps understanding resources in DLs

# Selected Modules for Scholars (4 of 7)

- 4-b:Metadata, cataloging, metadata markup, metadata harvesting
  - Helps scholars build DLs for specific domains
- 4-d:Subject description, vocabulary control, thesauri, terminologies
  - Subject description - useful when choosing keywords for ETDs
  - Vocabulary control helps organize a large number of digital objects

# Selected Modules for Scholars (5 of 7)

## ■ 5-c: Identifiers, handles, DOI, PURL

- Learn to preserve long-term access to digital objects (e.g., ETDs) in the network
- Learn about Superimposed Information technology [8] and present only the relevant portion of a paper

## ■ 6-b: Search strategy, info seeking behavior, user modeling

- Helps find relevant information by applying the various search strategies
- Study of info seeking behavior and user modeling helps design more effective and user-friendly DLs

# Selected Modules for Scholars (6 of 7)

- 7-e: Web publishing (e.g., wiki, RSS, Moodle, etc.)
  - Understand online tools with functions such as:
    - Forming groups with the same interests
    - Sharing of resources among the group members
    - Rating/commenting on the resources for quality control



# Selected Modules for Scholars (7 of 7)

- 9-b:DL case studies
  - Helps develop DLs for specific domains
- 9-e:Legal issues (e.g., copyright, privacy)
  - Helps avoid violation of the copyright laws for DL resources and the privacy of the member profile information in DLs

# Outline

1. DL educational module framework
2. Selected modules for scholars
3. Selected modules & mapping to ETD guide
4. Selected modules & their resources
5. Summary
6. Future work

# Selected Modules & Mapping to ETD Guide (1 of 4)

- Benefit: **save time**

- Read the relevant sections in the mapping table instead of browsing the whole ETD guide

- Modules with the **largest num.** of ETD guide sections

- 3-d: Document and e-publishing/presentation markup (27 sections)
- 4-b: Metadata, cataloging, metadata markup, metadata harvesting (12 sections)

➔ **Fundamental functions of an ETD system** (prepare papers and metadata)

# Selected Modules & Mapping to ETD Guide (2 of 4)

- Modules with the **smallest num.** of ETD guide sections
  - 3-b:Digitization (0 sections)
    - Possible reasons
      - ➔ETDs are created digitally nowadays
      - ➔No need to digitize paper-based theses and dissertations. Therefore, no reason to include sections on digitization

# Selected Modules & Mapping to ETD Guide (3 of 4) – Software Tools

- **Word processing tools** explained in the guide
  - MS Word, LaTeX, WordPerfect, MS Office 2000, Framemaker
  - As these tools are upgraded, related ETD guide sections should be updated, too
  - New sections are needed for new tools (e.g., Google docs & spreadsheets at <http://docs.google.com/>)

# Selected Modules & Mapping to ETD Guide (4 of 4) (selected portion)

<p>7-e:Web publishing (e.g., wiki, RSS, Moodle, etc.)</p>	<p>4.2.7 Packaging (page 289) 4.3.4.1 Packaged solutions (page 321) 4.3.4.1.1 DiTeD and DIENST (page 322) 4.3.4.1.2 ADT (page 328) 4.3.4.1.3 Cybertheses (page 331) 4.3.4.1.4 VT DB and other tools (page 333) 4.3.4.1.5 Library automation: OPACs, VTLS software (page 336)</p>
<p>9-b:DL case studies</p>	<p>5.1 Training the trainers: initiatives to support electronic theses and dissertation projects in Latin America (page 373) 5.3.1 Initiatives &amp; projects (page 388) 5.4 Creating an online database of problem solving solutions (page 397) 5.5 Help develop a broad local team, Australian digital theses program(page 399) 5.6 Standards, cooperation and collaboration, Australian digital theses program (page 400) 5.7.1 Developing centers of expertise, Australian digital theses program (page 405)</p>
<p>9-e:Legal issues (e.g., copyright)</p>	<p>2.3.1 Intellectual property rights (page 71) 3.2.8 Protecting intellectual property / how to deal with plagiarism (page 226) Plagiarism (page 93, no section number) 4.2.6 Encryption and watermarking (page 287)</p>

# Outline

1. DL educational module framework
2. Selected modules for scholars
3. Selected modules & mapping to ETD guide
4. Selected modules & their resources
5. Summary
6. Future work

# Selected Modules and Their Resources

## (Only two rows chosen from the table)

- Full table is at [curric.dlib.vt.edu](http://curric.dlib.vt.edu) -

1-a(10-c):Conceptual frameworks, theories, definitions	<ul style="list-style-type: none"><li>• Alexa T. McCray and Marie E. Gallagher, "Principles for Digital Library Development," Communications of the ACM, Volume 44, Issue 5 (May 2001). (ACM DL)</li><li>• Atkinson, Ross. (1996). Library functions, scholarly communication, and the foundation of the digital library: Laying claim to the control zone. Library quarterly 66 (3), pp. 239-65.</li><li>• Fox, Edward A., &amp; Marchionini, Gary (Eds.). (2001). Digital libraries. Communications of the ACM, 44(5).</li></ul>
2-c(8-c):File formats, transformation, migration	<ul style="list-style-type: none"><li>• Witten, I.H., &amp; Bainbridge, D. (2003). How to build a digital library. Morgan Kaufmann. (pp163-216)</li><li>• Lawrence, Gregory W., Kehoe, William R., Rieger, Oya Y., Walters, William H., &amp; Kenney, Anne R. (2000). Risk management of digital information: A file format approach. Washington, DC: CLIR.</li><li>• Carl Fleischhauer (1998) Digital Formats for Content Reproductions. Library of Congress. <a href="http://memory.loc.gov/ammem/formats.html">http://memory.loc.gov/ammem/formats.html</a></li></ul>



# Outline

1. DL educational module framework
2. Selected modules for scholars
3. Selected modules & mapping to ETD guide
4. Selected modules & their resources
5. Summary
6. Future work

# Summary

- We provided our latest 10-module framework which covers most of the topics in DLs and their surrounding environment. This framework is our ground to build DL educational lesson modules for students and scholars.
- From those modules, we chose eleven modules which we think the most relevant to scholars' research endeavors. Then those eleven modules' descriptions, their mapping to ETD guide sections and their corresponding readings from our syllabi reading collection were presented in the Tables which will be available on our project web site.
- It is our hope that the scholars who are not familiar working with ETDs and the ND LTD system and want to find the relevant ETDs for their research and submit their ETDs to the system will get benefit from our paper.



# Outline

1. DL educational module framework
2. Selected modules for scholars
3. Selected modules & mapping to ETD guide
4. Selected modules & their resources
5. Summary
6. **Future work**

# Future Work (1 of 2)

- Module-ETD guide mapping table will be added to the ETD guide

Link to the module framework and descriptions

Link to the corresponding ETD guide section page

DL Module	ETD guide section
1-a(10-c): Conceptual frameworks, theories, definitions	1. Introduction: Purpose and scope of this document (page 10) 1.1. What are ETDs? (page 12) 1.1.1. ETDs as a new genre of documents (page 14)
2-c(8-c): File formats, transformation, migration	3.2.4 Preparing PDF document (page 187) 3.2.4.2 PDF from LaTeX (page 188) 3.2.5 Preparing for conversion to SGML/XML (page 195) 3.2.5.1 Preparing for conversion to SGML/XML from MS Word (page 200) 3.2.5.4 Preparing for conversion to SGML/XML in LaTeX (page 213) 4.2.2 Page description languages (PostScript and PDF) (page 248) 4.2.3.2 DTDs for ETDs (page 256) 4.2.3.4 Conversions from Word to SGML/XML (page 265) 4.2.3.4.2 Conversions to SGML/XML from LaTeX (page 276) Rendering style-sheets (page 278) 6.3 Managing technology changes (page 409)
3-b: Digitization	No corresponding sections
3-d: Document and e-publishing/presentation markup	3.2.1 Overview: writing with word processors and structured editors (page 155) 3.2.2 Writing in word processing systems (page 157) 3.2.2.1 Microsoft Word and Office 2000 (page 158)

# Future Work (2 of 2)

- Module-Resources table will be added to the ETD guide


Link to the module framework and descriptions

Link to the original document page or PDF file download page

1-a(10-c): Conceptual frameworks, theories, definitions	<ul style="list-style-type: none"><li>• Alexa T. McCray and Marie E. Gallagher, "<u>Principles for Digital Library Development</u>," <u>Communications of the ACM</u>, Volume 44, Issue 5 (May 2001). (ACM DL)</li><li>• Atkinson, Ross. (1996). <u>Library functions, scholarly communication, and the foundation of the digital library: Laying claim to the control zone</u>. <u>Library quarterly</u> 66 (3), pp. 239-65.</li><li>• Fox, Edward A., &amp; Marchionini, Gary (Eds.). (2001). <u>Digital libraries</u>. <u>Communications of the ACM</u>, 44(5).</li></ul>
2-c(8-c) File formats, transformation, migration	<ul style="list-style-type: none"><li>• Witten, I.H., &amp; Bainbridge, D. (2003). <u>How to build a digital library</u>. Morgan Kaufmann. (pp163-216)</li><li>• Lawrence, Gregory W., Kehoe, William R., Rieger, Oya Y., Walters, William H., &amp; Kenney, Anne R. (2000). <u>File management of digital information: A file format approach</u>. Washington, DC: CLIR.</li><li>• Carl Fleischhauer (1998) <u>Digital Formats for Content Reproductions</u>. Library of Congress. <a href="http://memory.loc.gov/ammem/formats.html">http://memory.loc.gov/ammem/formats.html</a></li></ul>
3-b: Digitization	<ul style="list-style-type: none"><li>• Chowdhury, G.G., &amp; Chowdhury, Sudatta. (2003). Chapter 6 ? Digitization. (pp. 103-119). In <u>Introduction to digital libraries</u>.</li></ul>

# References

- [1] M. A. Gonçalves, “Streams, Structures, Spaces, Scenarios, and Societies (5S): A Formal Digital Library Framework and Its Applications,” Computer Science Doctoral Dissertation. Blacksburg, VA: Virginia Tech, 2004, 161 pages. <http://scholar.lib.vt.edu/theses/available/etd-12052004-135923/unrestricted/MarcosDissertation.pdf>
- [2] M. A. Gonçalves and E. A. Fox, “5SL - A Language for Declarative Specification and Generation of Digital Libraries,” in Proc. JCDL’2002, Second ACM / IEEE-CS Joint Conference on Digital Libraries, July 14-18, G. Marchionini, Ed. Portland, Oregon, USA: ACM, 2002, pp. 263-272.
- [3] M. Gonçalves, E. Fox, L. Watson, and N. Kipp, “Streams, Structures, Spaces, Scenarios, Societies (5S): A Formal Model for Digital Libraries,” ACM Transactions on Information Systems, vol. 22, pp. 270-312, 2004.
- [4] Fox, E. A., Wildemuth, B. M., Pomerantz, J., and Yang, S. Digital Library Curriculum Development: Enhancing Education and Comprehension of NDLTD, The 9th International Symposium on Electronic Theses and Dissertations (ETD 2006) at [http://www6.bibl.ulaval.ca:8080/etd2006/pages/papers/SP11\\_Ed\\_Fox.ppt](http://www6.bibl.ulaval.ca:8080/etd2006/pages/papers/SP11_Ed_Fox.ppt)

- 
- [5] Pomerantz, J., Wildemuth, B. M., Yang, S., and Fox, E. A. 2006. Curriculum development for digital libraries. In Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries (Chapel Hill, NC, USA, June 11 - 15, 2006). JCDL '06. ACM Press, New York, NY, 175-184. DOI= <http://doi.acm.org/10.1145/1141753.1141787>
- [6] Pomerantz, J., Oh, S., Wildemuth, B., Yang, S. and Fox, E. A. (2007). Digital Library Education in Computer Science Programs. The 7th ACM/IEEE-CS Joint Conference on Digital Libraries (Vancouver, British Columbia, Canada, June 18-23, 2007). JCDL '07.
- [7] Pomerantz, J., Oh, S., Yang, S., Fox, E. A., & Wildemuth, B. (2006). The Core: Digital Library Education in Library and Information Science Programs. D-Lib Magazine, 12(11).
- [8] The Superimposed Information Project at Virginia Tech, Blacksburg, VA, 2006. URL: <http://si.dlib.vt.edu/>



Thank you !