Further Development of a Digital Library Curriculum: Evaluation Approaches and New Tools

Seungwon Yang[1], Barbara M. Wildemuth[2], Seonho Kim[1], Uma Murthy[1], Jeffrey P. Pomerantz[2], Sanghee Oh[2], and Edward A. Fox[1]

[1] Department of Computer Science, Virginia Tech, Blacksburg, VA 24061 U.S.A.
{seungwon,shk,umurthy,fox}@vt.edu
Tel.: +1 540-231-5113

[2] School of Information and Library Science, University of North Carolina – Chapel Hill, Chapel Hill, NC 27599-3360 U.S.A.
wildem@ils.unc.edu, {jpom,shoh}@email.unc.edu
Tel.: +1 919-962-8366

Abstract. This paper is a follow-up to our ICADL 2006 paper, reporting on our progress over the past year in developing a digital library curriculum. It presents and describes the current curriculum framework, which now includes ten modules and 41 sub-modules. It provides an overview of the curriculum development lifecycle, and our progress through that lifecycle. In particular, it reports on our evaluation of the modules that have been drafted. It concludes with a description of two new technologies – Superimposed Information (SI) to help resource presentation in a module and Visual User model Data Mining (VUDM) to help long-term module upgrade by visualizing the user community and its trends.

Keywords: Digital library, curriculum, education, evaluation, superimposed information, community visualization.

1 Introduction

Our Digital Library (DL) Curriculum Development Project[1] [3] is now in its second year (of three years of funding). Since describing the project at ICADL 2006 in Kyoto, Japan, our curriculum framework has continued to evolve, based on analyses of the literature and course syllabi in information and library science and in computer science. In addition, we have developed draft versions of seven sub-modules, and have conducted a pilot test of our formative evaluation procedures.

In section 2, we show the recently-updated DL curriculum framework, which has ten core modules, which in turn have 41 sub-modules. Section 3 presents our curriculum development lifecycle, and discusses our progress through that lifecycle. In particular, it describes our plans for expert review of draft modules, already begun at a meeting held during the Joint Conference on Digital Libraries, June 2007, and our

[1] Funded by US NSF Grants IIS-0535057 to VT and IIS-0535060 to UNC-CH for 2006-2008.

D.H.-L. Goh et al. (Eds.): ICADL 2007, LNCS 4822, pp. 434–443, 2007.

plans for field testing of the revised modules. In section 4, we introduce two software tools – our Superimposed Information (SI) tool [16] and our user community/trend visualization tool, i.e., Visual User model Data Mining (VUDM) [10]. They will be used to enhance the presentation of papers in a module (SI tool) and to help long-term upgrading of the DL modules by visualizing user trends over time (VUDM tool).

2 Evolution of the DL Curriculum Framework

New technologies are emerging in every area of computer science such as databases, human-computer interaction, web-based technologies, multimedia, hypermedia, search algorithms, security, etc. Considering that a digital library (DL) integrates most of those technologies under a unified system, the topics in the DL area are changing as well.

The model presented in Fig. 1 is the fourth and current version of a DL curriculum framework developed for this project.[2] Initially, our work was based on four sources. The first was an analysis of the Computing Curriculum 2001 (CC2001), recommendations for undergraduate curricula and course content in computer science. CC2001 recommends that Digital Libraries be included as one component of education on Information Management. Second, we built our initial framework on the 5S theoretical framework [5,6,7]. This framework specifies streams, structures, spaces, scenarios, and societies as the core components of digital libraries and, thus, necessary components of a curriculum on digital libraries. Third, our framework was influenced by the results of a survey of digital librarians, concerning the skills and knowledge required to manage a digital library [2]. Fourth, we based our initial framework on our own experience in teaching courses on digital libraries.

Since then, the project team has been refining the curriculum framework through a series of analyses and classification tasks. First, we analyzed the recent literature in digital libraries [14], by classifying the papers presented at JCDL and its predecessor conferences or published in *D-Lib Magazine* over the past decade. The core modules of our curriculum framework were used as the basis for the classification of 1064 papers; in addition, when a paper could not be easily classified, it initiated a discussion of possible revision of our framework. Next, we similarly analyzed syllabi from 40 digital library courses offered by 29 U.S. programs in information and library science [15]. Finally, we analyzed syllabi from digital libraries courses offered in computer science (CS) programs in the U.S. [13]. Only five courses focusing on digital libraries were identified from an examination of the websites of 296 CS programs. The readings from these five courses were added to our analysis. During each analysis, there were two types of gaps between the analytical findings and the curriculum framework that were investigated. First, if very few papers were classified into a module in the framework, we discussed whether the topics could be merged together or one of them should be removed from the framework. Second, if an unusually high number of readings were classified into the same module, we discussed the possibility of splitting that module.

[2] Earlier versions are available on the project's Module Development web page, http://curric.dlib.vt.edu/modDev/modDev.html

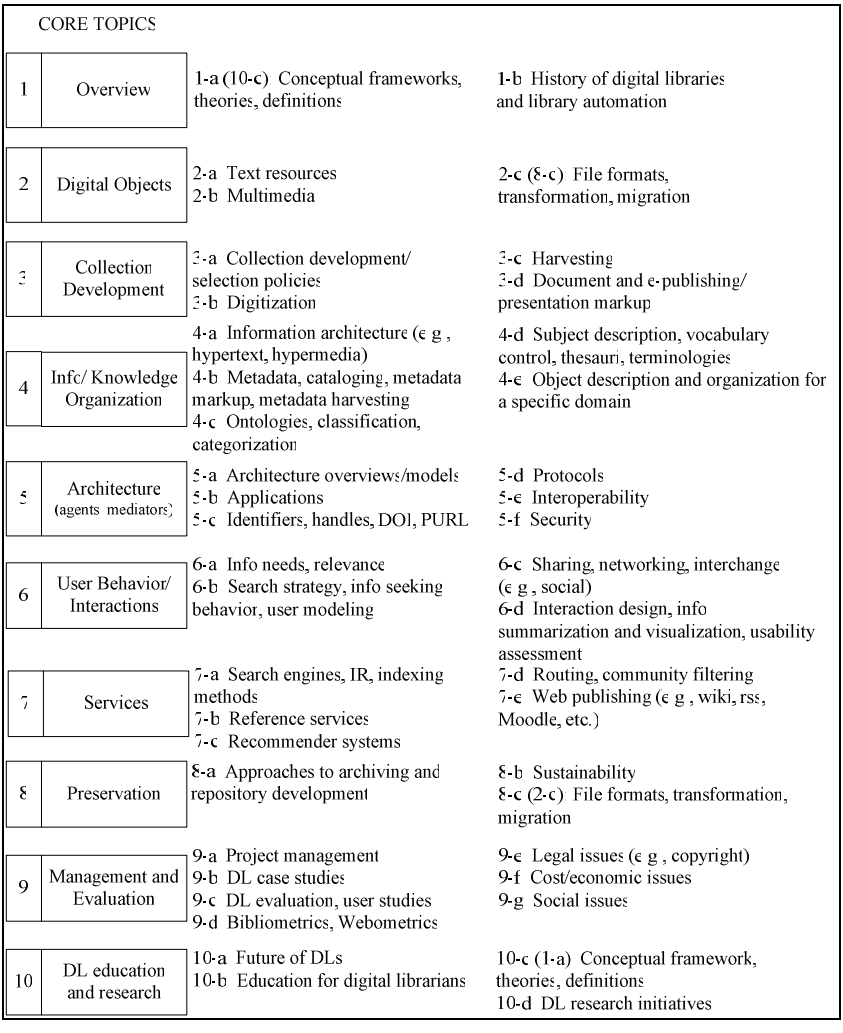

Fig. 1. DL curriculum framework

Based on these discussions, several changes were made in the curriculum framework, to bring it to the form found in Fig. 1. First, the order of the two core modules, '2: Digital object' and '3: Collection development', were reversed to be more natural (moving from considerations of a single digital object to a group of them). Second, two new sub-modules were introduced: '1-b: History of digital library and library automation' and '4-a: Information architecture (e.g., hypertext, hypermedia)'.

While we believe that the current framework is stable enough to serve as the basis for development of draft modules, it is likely that it will continue to evolve. In particular, it is important to remember that this framework's development was based on present and past teaching practices and published resources. Therefore, the current

framework does not yet include emerging topics in the DL field (except in 10-a and 10-d). As evidence of their importance is identified, however, those topics will be incorporated into future versions of the module framework.

3 Digital Library Curriculum Development Lifecycle

Fig. 2 shows the development process in a spiral lifecycle, as suggested by Boehm's [1] model of the system development lifecycle. Currently, we are focused on three stages: 'Design modules', 'Evaluate via inspection', and 'Feedback on strengths & weaknesses.' In Fig. 2, these three are marked in yellow.

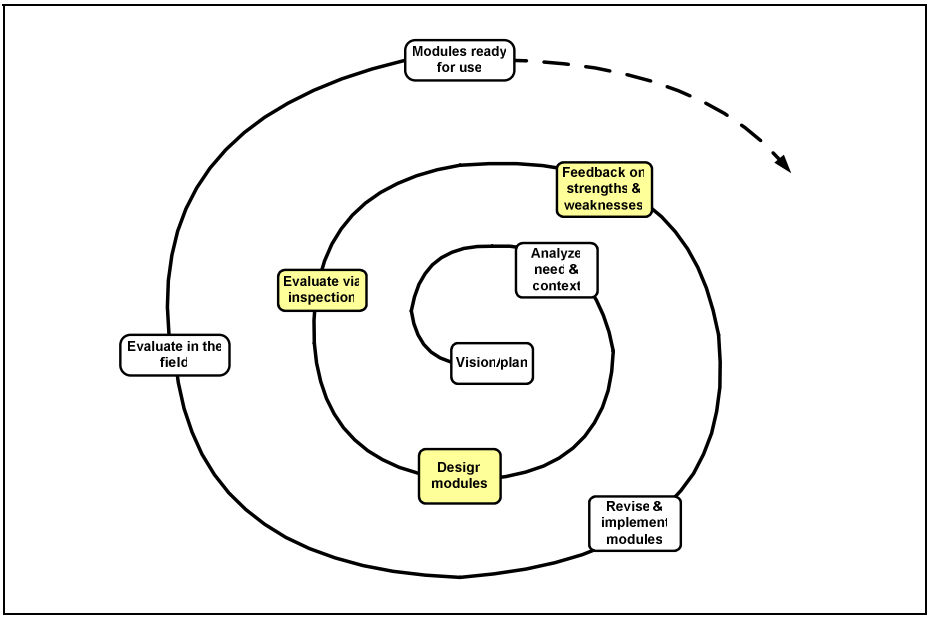

Fig. 2. DL curriculum development lifecycle

For the 'Design modules' step, in this iteration through the cycle, we developed seven draft modules. These modules were selected to cover some of the high-priority areas of the curriculum, as suggested by earlier analyses [2,13,14]. They are:

- 1-b: History of digital libraries and library automation. The origin of the DL research agenda, DLI, DLI-2, NSDL, and the origin of other long-term DL projects still extant.
- 5-a: Architecture overviews/models. Different types of DL architectures and models such as operational, technical, systems, component-based, federated, distributed, and service-oriented architectures.
- 5-b: Applications. Different types of DL systems, such as repository-based systems (e.g., DSpace, EPrints, Greenstone, FEDORA), metadata-based

systems (e.g., CiteSeer), and component-based systems (e.g., Open Digital Libraries – ODL – and WS-ODL).

- 6-a: Information needs, relevance. Aspects of a user's experience of an information need and how that experience might affect the user's interactions with the DL; relevance judgments and their relationship to the user's information needs.
- 6-b: Search strategy, information seeking behavior, user modeling. The fundamental concepts, definitions, and theoretical models of online information seeking behaviors, as they apply to digital libraries; user behaviors that have been identified in empirical studies of digital libraries.
- 7-b: Reference services. Services for meeting different types of user information needs addressed to digital libraries, including human-mediated reference, automated information retrieval (IR), and question answering (QA) services.
- 9-c: DL evaluation, user studies. Methods for evaluating the outcomes, impacts, or benefits of a digital library, including cost/benefit analyses.

Each module has the same structure following the module template [3]. Key components of each module include its learning objectives, the prerequisite knowledge required and its relationship to other modules, the body of knowledge to be covered, assignments and learning activities, and readings and other resources supporting the module.

The seven modules listed above are now ready for the next step in the curriculum development lifecycle: evaluation via inspection by experts. To pilot test our procedures for these evaluations, the project team convened a meeting during the 2007 Joint Conference on Digital Libraries (JCDL). It included members of our Advisory Board, participants in the JCDL Doctoral Consortium, and other members of the community with particular interest in curriculum evaluation. We first provided an overview of our progress in developing modules, and reviewed the curriculum development lifecycle. The participants then worked in pairs to evaluate/review a module (each pair selected a module of particular interest to them). They were guided in their evaluation by an evaluation form provided by the project team (see Fig. 3), and developed based on the work of Diamond [4], Grunert [8], and Wiggins and McTighe [17].

The final portion of the meeting was spent discussing the evaluation process itself, and how it should be implemented more widely. We received several suggestions on the module development and evaluation process. The first is to prioritize the modules, so that the most critical modules are developed first. To prioritize them, we will conduct a study of the distribution of the topics appearing in the table of contents of various DL textbooks, in the articles of DL-related magazines and journals, and in the DL class syllabi. The second suggestion is to define the scope of each module and to specify the dependencies and relationships among the modules more consistently. We will undertake this effort in the near future and will conclude it before developing an additional set of modules. Third, the group encouraged us to include more international scholars and doctoral students in the development process, so that the modules would be usable outside the U.S.

[3] http://curric.dlib.vt.edu/DLcurric/moduleTemplate.html

1. **Objectives:**
   Guiding question: Are the objectives appropriate for the topic?
   Specific questions:
   - Are the objectives observable?
   - Will students be able to achieve the objectives, given the content in the body of knowledge?
2. **Body of knowledge:**
   Guiding question: Does the module address all areas of the topic that need to be addressed?
   Specific questions:
   - Will the body of knowledge enable students to achieve the objectives?
   - Are there any topics that you think are critical to add to the body of knowledge?
   - Are there any topics that you would remove from the body of knowledge?
3. **Readings:**
   Guiding question: Are the readings the best and most appropriate for the topic?
   Specific questions:
   - Are there any readings that you think are critical to add to the list?
   - Are there any readings on the list that you would remove?
4. **Learning Activities:**
   Guiding question: Are the activities appropriate for the topic?
   Specific questions:
   - Will students be able to accomplish the activities, given the content in the body of knowledge?
   - Will the activities enable students to achieve the objectives?
   - Can you suggest any other learning activities that may be appropriate for this module?
5. **Level of Effort and Prerequisites:**
   Guiding question: Is it feasible to teach the module as it is currently constructed?
   Specific questions:
   - Is the level of effort required in class appropriate to the scope of the body of knowledge? Prior to class?
   - Is the prerequisite knowledge required sufficient for students to comprehend the body of knowledge?

Fig. 3. DL module formative evaluation form

As we continue the evaluation process, evaluators will be identified through individual nominations, review of the membership of the American Society for Information Science & Technology (ASIST) Special Interest Group on Digital Libraries (SIG DL), review of the attendance list for JCDL 2007, and other sources. To support the involvement in the evaluation process of a wide variety of experts in DLs, we will establish a password-protected wiki. Evaluators will be invited (or may volunteer) to evaluate a particular module of interest to them, and will be supplied with the wiki password. The current version of the module will be available for their review; a separate discussion page for each module will be established, so that people may comment on the module. As changes to the module are suggested and consensus

is reached about the need for those changes, a revised version of the module will be provided and the wiki-based discussion will be edited to contain only those issues not yet resolved. In this way, any number of evaluators may be involved in discussing and suggesting improvements for each module. The project team will closely monitor these discussions to provide support and to finalize the module draft when the discussion has reached a conclusion.

When each module has been evaluated, the feedback received will be incorporated into its design. The next step is to field test the module. Again, volunteers will be recruited to implement particular modules in their regular classes. These classes may include those focused on digital libraries, or other classes in which a particular module would be useful. We will check the implementation of each module by tracking which portions of the module were used as proposed and which were modified, and we will capture the actual assignments completed in connection with the module. At the completion of the module in each class, we will interview the instructor and survey the students. The instructor will be asked the same types of questions that are included on the expert review evaluation form (see Fig. 3); the students will be asked about the ways in which the module (its content, readings, and assignments) affected their learning. The results from these evaluations will be available via our project website, as well as being published in more formal venues, so that potential users of each module can adapt the module as appropriate for a particular situation. At that point, we will have reached the end of our curriculum development lifecycle, with 'Modules ready for use'.

4 Future Work

We plan to incorporate two new technologies. The Superimposed Information (SI) technology will be used to enhance module resource presentation by displaying only the relevant portion of a resource. Therefore, it will help students save time studying the module resources assigned to them. Once all the modules are developed and deployed, Visual User model Data Mining tool (VUDM) will be used to visualize the module user groups and analyze the usage trend over time for further module update.

4.1 Superimposed Information (SI) Technology

In many educational tasks, there is a need to be specific about a reference – to operate at a finer level of granularity than a complete document. For example, a user may want to work with a definition in a paper, a section in a book, or part of an image or a short clip in an audio/video document. Current annotation and knowledge management tools support this functionality to some extent. However, they provide limited support for working with information at sub-document granularity across heterogeneous formats. Using SI technology, users may lay new information over existing or *base information*, typically to highlight, annotate, elaborate, select, collect, organize, connect, or reuse information elements. This functionality enables the user to work with information at sub-document granularity *in situ* (in its original form and context) [11]. These tools employ *marks*, which are references to selected regions within base information (of text or multimedia content) [12]. Fig. 4 shows how a

mark may be used as part of the list of resources for the DL application module, and the resolved mark highlighting the selection that describes DSpace. In addition, this mark may be used in any place where a URI[4] may be used, such as in a module web page, as part of a concept map[5] about a module, etc. For details on work done on SI by the authors and their collaborators, please refer to [16].

We believe that using superimposed tools in the DL curriculum project has advantages. They allow users to select and work with sub-document information, while simultaneously retaining links to the original information context. This can be beneficial in providing greater specificity of reference in multiple ways:

- In module development: Using superimposed tools to describe resources in modules will help in focusing on part of a resource while still providing links to the complete documents.
- In module usage: Students and instructors can customize/personalize their use of modules in various assignments and projects by being specific about a part of a resource.

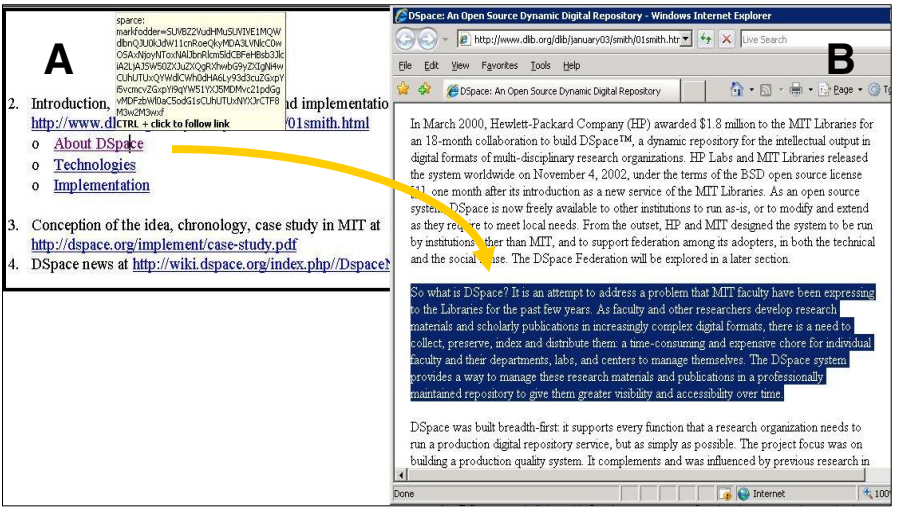

Fig. 4. A) Mark (represented as a URI) used in a module; B) Mark highlighting the desired selection that describes DSpace in an article

4.2 User Community/Trend Visualization Tool – VUDM

Visual User model Data Mining tool, VUDM [10], was designed to visualize users, user communities, and usage trends of complex information system, e.g., to analyze the DL curriculum module usage. VUDM visualizes user communities based on the

[4] Short for ***Uniform Resource Identifier,*** the generic term for all types of names and addresses that refers to objects on the World Wide Web.

[5] A concept map is a graphical knowledge representation tool, where the nodes represent concepts and the links represent relationships between concepts.

long-term history of usages of each user, instead of an explicitly-entered user profile. A web-based curriculum module server will be built to collect the usage history of each curriculum module for all users.

Fig. 5 illustrates VUDM visualizing the change in usage trends of curriculum modules for three consecutive weeks. Module number is displayed for each user group. By observing the spirals and their module number each week, we can see the changes of the user groups' interests (e.g., a group might become bigger, smaller or disappear). This information will be useful for understanding the evolution of DL curriculum modules. Further, VUDM is able to visualize the trends of positive or negative rankings entered by module users through the distribution web site. This trend indicates controversial, erroneous, and out-dated modules, and is useful for module upgrade if analyzed along with users' comments from the online forum.

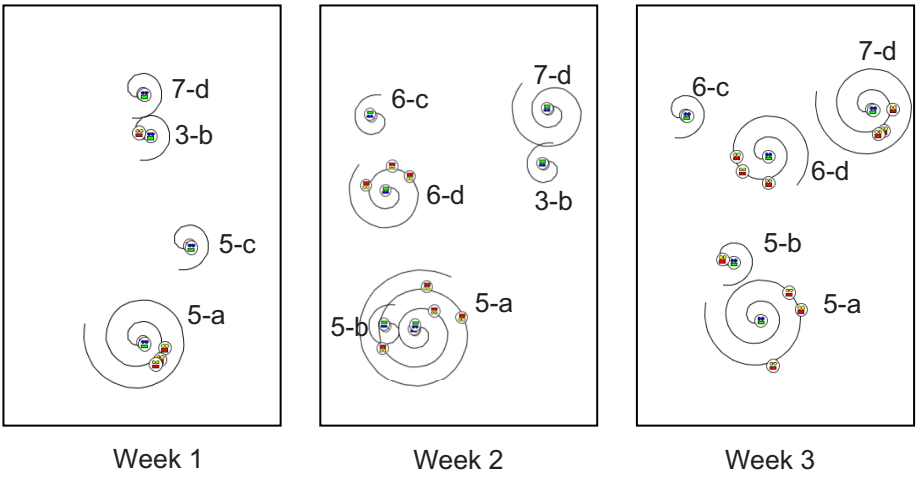

Fig. 5. Visualization of DL module usage trends for three weeks

5 Conclusion

The collaborative Virginia Tech - University of North Carolina DL curriculum development project is in the second of its three years. We have developed several curricular modules and have pilot tested our preliminary evaluation methods. In this paper, we presented the upgraded DL module framework, the curriculum development lifecycle, the draft module evaluation form, and two technologies that will support the presentation of the resources in the module (SI tool) and will help the project group to understand the module users' needs by visualizing the user community topics and trend changes (VUDM tool). We hope that the international DL community will become actively involved in this interdisciplinary effort, and that its results will improve the education of DL professionals.

References

1. Boehm, B.W.: A spiral model of software development and enhancement. IEEE Computer 21, 5, 61–72 (1988)
2. Choi, Y., Rasmussen, E.: What is needed to educate future digital librarians: A study of current practice and staffing patterns in academic and research libraries. D-Lib Magazine 12, 9
3. DL Curriculum Project Group: Collaborative Research: Curriculum Development for Digital Library Education (2006), http://curric.dlib.vt.edu/
4. Diamond, R.M.: Designing & Assessing Courses & Curricula: A Practical Guide. Jossey-Bass, San Francisco (1998)
5. Gonçalves, M.A.: Streams, Structures, Spaces, Scenarios, and Societies (5S): A Formal Digital Library Framework and Its Applications. Ph.D. dissertation. Computer Science Dept., Virginia Tech, Blacksburg, VA (2004), http://scholar.lib.vt.edu/theses/available/etd-12052004-135923/unrestricted/MarcosDissertation.pdf
6. Gonçalves, M.A., Fox, E.A.: 5SL - A language for declarative specification and generation of digital libraries. In: Proceedings of JCDL 2002, Second ACM / IEEE-CS Joint Conference on Digital Libraries, pp. 263–272 (2002)
7. Gonçalves, M., Fox, E., Watson, L., Kipp, N.: Streams, structures, spaces, scenarios, societies (5S): A formal model for digital libraries. ACM Transactions on Information Systems 22, 270–312 (2004)
8. Grunert, J.: The Course Syllabus. Anker Publishing Company, Inc., Bolton, MA (1997)
9. Gustafson, K.L., Branch, R.M.: Survey of Instructional Development Models, 3rd edn. ERIC Clearinghouse on Information Resources, Syracuse University, Syracuse, NY (1997)
10. Kim, S., Lele, S., Ramalingam, S., Fox, E.A.: Visualizing user communities and usage trends of digital libraries based on user tracking information. In: Sugimoto, S., Hunter, J., Rauber, A., Morishima, A. (eds.) ICADL 2006. LNCS, vol. 4312, pp. 111–120. Springer, Heidelberg (2006)
11. Maier, D., Delcambre, L.: Superimposed information for the Internet. In: WebDB Workshop, Philadelphia, PA, June 3-4, pp. 1–9 (1999)
12. Murthy, S., Maier, D., Delcambre, L., Bowers, S.: Putting integrated information into context: Superimposing conceptual models with SPARCE. In: First Asia-Pacific Conference of Conceptual Modeling, Denedin, New Zealand, pp. 71–80 (2004)
13. Pomerantz, J., Oh, S., Wildemuth, B.M., Yang, S., Fox, E.A.: Digital library education in computer science programs. In: Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries, Vancouver, pp. 177–178 (2007)
14. Pomerantz, J., Oh, S., Yang, S., Fox, E. A., Wildemuth, B.: The core: Digital library education in library and information science programs. D-Lib Magazine, 12, 11 (2006), http://dlib.org/dlib/november06/pomerantz/11pomerantz.html
15. Pomerantz, J., Wildemuth, B., Fox, E. A., Yang, S.: Curriculum development for digital libraries. In: Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries, pp. 175–184 (2006)
16. Murthy, U., Fox, E.A.: Sidewalk at Virginia Tech. 2006. Digital Library Research Laboratory, Computer Science Dept., Virginia Tech, Blacksburg, VA, http://si.dlib.vt.edu/
17. Wiggins, G., McTighe, J.: Understanding by Design, 2nd edn. Association for Supervision and Curriculum Development, Alexandria, VA (2005)