

Workshop On Integrating Digital Library Content with Computational Tools and Services

A Full Day Workshop

19 June 2009

ACM/IEEE Joint Conference on Digital Libraries (JCDL 2009)
Austin, Texas, USA

J. Stephen Downie, Workshop Chair
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign

jdownie@illinois.edu

Over the past ten years, the development, deployment and use of “first generation” digital libraries (DL) have matured into a stable use paradigm: the browsing, searching and then retrieving of digital materials. With recent strides being made in the areas of data mining, high performance computing, semantic web, linked-data, and web-services, etc. exciting new opportunities are arising to create “second generation digital libraries” (SGDL) by extending the standard DL use paradigm to include the *analysis* of the retrieved materials in a tightly integrated manner. It is the purpose of this workshop to bring together all those that are interested in creating SGDL systems by making this analytic extension to the DL use paradigm a reality.

The workshop organizers are currently involved in several SDGL technology projects that deal with both humanities and scientific data sets:

1. Networked Environment for Music Analysis (NEMA; <http://nema.lis.uiuc.edu>)
2. Metadata Data Offer New Knowledge (MONK; <http://monkproject.org>)
3. MyExperiment (<http://myexperiment.org>)
4. Networked Environmental Sonic-Toolkits for Exploratory Research (NESTER; <http://nester.lis.uiuc.edu>)
5. The Software Environment for the Advancement of Scholarly Research (SEASR; <http://seasr.org>) and its Meandre infrastructure system (<http://seasr.org/meandre>).

In order to encourage the exchange of technologies, experiences and ideas, the workshop is soliciting participants drawn from the following communities:

1. CONTENT SPECIALISTS
 - a. **Content Scholarship:** Researchers, scholars and educators with interest in more fully engaging with and more fully exploiting the content of digital libraries and similar repositories of digital materials
 - b. **Content Provision:** Digital librarians, digital archivists and data curators, etc. who want to better serve their users by providing extended mechanisms by which users can select, manipulate and then perform sophisticated analyses on the selected materials.
2. DEVELOPMENT SPECIALISTS
 - a. **Digital Library Development:** Digital library software developers and administrators, etc. who are engaged at the coding level with the creation, modification and/or administration of digital library and digital repository software.
 - b. **Tool and Service Development:** Computer scientists, semantic web/linked-data experts, web-services developers, data mining researchers, computation musicologists, and digital humanists, etc. who are developing analytic tools that should be better integrated into digital library and digital repository installations.

Objectives:

1. Foster the development of methods and technologies to create SGDL systems by:
 - a. having those on content side inform those on the development side about real-world problems, needs, and opportunities
 - b. having those on the development side inform those on the content side about the strengths and weaknesses of the various approaches to creating SGDL systems
2. Foster a framework for future fruitful research into SGDLs by having participants explore consensus opinion on the establishment of research priorities, inter-disciplinary collaborations, evaluation standards, test collections, resource sharing, funding opportunities, communication channels, etc.

Approach:

1. **Integrate participants:** As stated in the objectives, we intend to bring together rather disparate groups that collectively have an interest in SGDLs. Throughout the course of the workshop we will pay particular attention to the effective communication between the groups. In short, we are striving for a rigorous and accurate discussion of the issues that at the same time minimizes jargon and specialist shorthand.
2. **Lecture/Demonstration/Discussion Format:** After the background material has been covered the primary mode of communication for the workshop will be a series of Lecture/Demonstration/Discussion presentations. In the Lecture presenters will explicate the motivations and strengths of a particular approach. Demonstrations (where applicable) of the approach will be made to make the Lecture more concrete. Discussion will be encouraged after each Lecture/Demonstration.
3. **Panelist Response:** In all likelihood, panelists and presenters will be one and the same (though not necessarily). However, the presenters will be asked to reconstituted themselves as members of the panel. The panelists will provide commentary in response to the various presentations. It is hoped that this phase of the workshop will lead to the identification of some common intellectual threads or points of debate.
4. **Constitutional Forum:** The last period of the workshop (say between the last coffee-break and the close of the workshop) will be the open-floor discussion (open to participants and presenters alike) concerning how to move forward. We will propose, based upon the discussion, the formation of some kind of more formal organization for those interested.

Call for Participation:

We invite the participation of all interested parties. You may participate at one or more level(s):

1. Presenter
2. Panelist
3. General Participant

Presentations and/or demonstrations are solicited that cover such topics as:

1. Prototype or deployed systems that have SGDL capabilities
2. Prototype or deployed analytic toolkits that could play a role in creating SGDL systems
3. Use case scenarios based upon actual or intended SGDL capabilities
4. Intellectual property issues surrounding content accessibility and exploitation
5. Intellectual property issues surrounding the integration of foreign tools into extant DL systems (i.e., license incompatibilities).
6. Interface, security and user management issues
7. Computational infrastructure issues (i.e., providing compute cycles for the analyses and storage of results)
8. Standards, either extant or proposed, that can help generalize the integration process across unique collections and tools

Formal papers are not required. However, we do want to have an extended abstracts 2-4 pages (JCDL format) for the demos and presentations along with appropriate presentation materials (copies of slides, charts, etc.) for inclusion in the workshop information package.

Submission requirements:

Whether you want to participate as a General Participant, Presenter and/or Panelist, we would like you to contact the Workshop Chair ASAP. We would like to have your name, background information, contact information, and the level at which wish to participate. For those wishing to present we would also appreciate brief abstract of your proposed lecture/demonstration along with any special equipment requirements. (NOTE: this does not constitute formal registration; it is merely a means by which we can better plan the workshop).

Submission Information:

Send all expressions of interest, proposals, and questions, to the Workshop Chair, J. Stephen Downie [jdownie@illinois.edu] with subject: SGDL Workshop.

Deadlines:

10 May 2009 for submitting proposals for lectures and/or demonstrations

10 June 2009 for submission of “camera ready” package materials.

Workshop Organizer(s):

Dr. J. Stephen Downie is an Associate Professor at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign (UIUC). He is Director of International Music Information Retrieval Systems Evaluation Laboratory (IMIRSEL). Professor Downie is Principal Investigator on the Networked Environment for Music Analysis (NEMA; <http://nema.lis.uiuc.edu>). He has been very active in the establishment of the Music Information Retrieval and Music Digital Library communities through his ongoing work with the ISMIR series of MIR Conferences as a member of the ISMIR steering committee.

Dr. David De Roure is a Professor of Computer Science in the School of Electronics and Computer Science at the University of Southampton, UK. A founding member of the School's Intelligence, Agents, Multimedia Group, he leads the e-Research activities and is Director of the Pervasive Systems Centre. His current research interest is the application of Semantic Web, Web 2.0 and workflow technologies to enable new research in multiple disciplines. He is a pioneer of the Semantic Grid and is closely involved in UK e-Science programme activities including the CombeChem and myGrid projects, OMII-UK and myExperiment. David has worked for many years with distributed information systems and distributed programming languages, and has been active in the Web and hypertext communities. He is a Scientific Advisory Council member of the Web Science Research Initiative and a Fellow of the British Computer Society

Contact information for Workshop Chair:

Dr. J. Stephen Downie, Associate Professor
Graduate School of Library and Information Science,
University of Illinois at Urbana-Champaign,
501 East Daniel St., Champaign, IL,
61820-6212
fax: (217) 244-3302
Email: jdownie@illinois.edu