Welcome to Champaign-Urbana and the nineteenth meeting of the ACM/IEEE Joint Conference on Digital Libraries (JCDL). This is the first JCDL to be held in Champaign-Urbana; but the cities have a rich history of gatherings for the study of information science. The tradition runs deep in the area notably the Allerton Park Institute conferences that were held for nearly thirty years (https://www.ideals.illinois.edu/handle/2142/420) by the University of Illinois Information School. Several active and stalwart members of the JCDL community may even recall the 1995 Allerton Institute on User Centered Digital Library Design and Evaluation. It is with this storied tradition of collective research purpose and common cause for information research in all its many interdisciplinary threads that we open the 2019 JCDL conference!

This year we accepted 26 of the 88 full papers submissions (29.5%) and 16 of the 66 short papers submitted (24.2%). The papers were evaluated through careful review and discussion, including reviews, metareviews and a virtual program committee meeting through teleconference over multiple days. The program chairs are grateful to the members of the program committee who volunteered to evaluate the papers and especially to those who gave their time to meet at times that due to time zones required very early or late sessions for some.

The program chairs also wish to recognize the work of the venue sub-chairs, and in particular to note the contribution of the Posters/Demonstrations chairs — Boots Cassel, Milena Dobreva, and Jerome McDonough. This year, there were 93 Posters/Demonstrations submissions; 52 posters and 8 demonstrations were selected for presentation at the conference.

JCDL 2019 continues the tradition of being a robust, international conference. The conference submissions included authors from 29 countries, with the United States in the lead followed closely by China and Germany. Members of the program committee add another four countries to those represented.

JCDL is pleased to have the opportunity to recognize our highest quality submissions through the Vannevar Bush Best Paper Award and the Best Student Paper Award. These awards are chosen from nominated full papers by selected members of the program committee and are based on the published versions of the nominated papers. In addition, an award for Best Poster/Demonstration will be given based on attendee votes during the Poster/Demonstration session.

We are excited to announce the names of the nominees for the best paper awards. In addition to the recognition at the conference, the nominees will be invited to submit an extended version of their paper for further review and consideration for a special issue of the International Journal on Digital Libraries. The nominees are:

- Sawood Alam, Michele Weigle, Michael Nelson, Fernando Melo, and Daniel Gomes. “MementoMap framework for flexible and adaptive web archive profiling.”
- George Buchanan and Dana McKay. “One way or another I’m gonna find ya: The influence of input mechanism on scrolling in complex digital collections.”
- Drahomira Herrmannova, Nancy Pontika, and Petr Knoth. “Do authors deposit on time? Tracking open access policy compliance.”
- Dattatreya Mohapatra, Abhishek Maiti, Sumit Bhatia, and Tanmoy Chakraborty. “Go wide, go deep. Quantifying the impact of scientific papers through influence dispersion trees.”
- Catherine C. Marshall and Frank Shipman. “The ownership and control of online photos and game data: Patterns, trends, and keeping pace with evolving circumstances.”
- Yasunobu Sumikawa, Adam Jatowt, Antoine Doucet, and Jean-Philippe Moreux. “Large scale analysis of semantic and temporal aspects in cultural heritage collection’s search.”

The recipients of the awards will be announced at the conference.

We have three excellent keynotes speakers who will address several of our community interests in digital innovation, digitization and preservation, and access to scholarship; they each bring new and valuable perspectives to our conference themes of “Curated Knowledge, “Connected People,” and “Extraordinary Results.”

- Patricia Hswe, Program Officer, Andrew W. Mellon Foundation
- Robert Sanderson, Enterprise Data Architect, J. Paul Getty Trust
- John Wilkin, University Librarian, University of Illinois.

We would like also to express our appreciation for individual members of the conference committee who have generously contributed their time and expertise to the JCDL enterprise, we could not have done it without the very able contributions of all the members of the conference committee, including the assistance of our meta-reviewers: Lillian Cassel, Sally Jo Cunningham, Nicola Ferro, Edward Fox, Dion Goh, Daqing He, Adam Jatowt, Min-Yen Kan, Martin Klein, Gary Marchionini, Catherine Marshall, Philipp Mayr, Ian Milligan, Michael Nelson, Erich Neuhold, Vivien Petras, Edie Rasmussen, Frank Shipman, Hussein Suleman, Herbert Van de Sompel, and Dan Wu — all who took extra time and effort to devote their expertise and shape a high-quality program for JCDL 2019.

Of course, an international conference such as this doesn’t just happen without “boots on the ground” so speak, and much groundwork was undertaken by the Local Planning Committee, so we would like to underscore the efforts of our local planning team here in Champaign, IL, including Emily Knox and Brent M. West for their excellent planning and project management work! We are very thankful for the expert assistance of our JCDL planning assistant from the University of Illinois I-School, Alaine Martaus, PhD. Finally, the proceedings you are enjoying took considerable coordination to come to fruition from Alaine, and our publications committee of Maria Bonn and Dan Wu.

On behalf of our generous sponsors and all our contributors, we welcome everyone to JCDL 2019!

Program Chairs
Suzie Allard, School of Information Sciences, University of Tennessee Knoxville
Richard Furuta, Department of Computer Sciences, Texas A&M University
Atsuyuki Morishima, Graduate School of Library, Information and Media Studies, University of Tsukuba

Conference Chairs
J. Stephen Downie, School of Information Sciences, University of Illinois at Urbana-Champaign
Jim Hahn, University Library, University of Illinois at Urbana-Champaign
Qinghua Zhu, School of Information Management, Nanjing University
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunday, June 2</td>
<td><strong>Tutorial (1)</strong> Introduction to Digital Libraries</td>
<td><strong>Illini Union - Room 314A</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Doctoral Consortium</strong></td>
<td><strong>Illini Union - Room 314B</strong></td>
</tr>
<tr>
<td>8:00am-12:00pm</td>
<td><strong>Lunch</strong></td>
<td></td>
</tr>
<tr>
<td>12:00 - 1:00pm</td>
<td><strong>Lunch</strong></td>
<td></td>
</tr>
<tr>
<td>1:00 - 5:00pm</td>
<td><strong>Tutorial (2)</strong> Strategies and Tools for Digital Repository Selection and Migration</td>
<td><strong>Illini Union - Room 210</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Tutorial (3)</strong> Building Digital Library Collections with Greenstone 3</td>
<td><strong>Illini Union - Room 407</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Tutorial (4)</strong> Preparing Code and Data for Computational Reproducibility</td>
<td><strong>Illini Union - Room 314A</strong></td>
</tr>
<tr>
<td>5:00 - 6:00pm</td>
<td><strong>Steering Committee Meeting</strong></td>
<td><strong>I-Hotel Quad Room</strong></td>
</tr>
<tr>
<td>6:00 - 8:00pm</td>
<td><strong>Welcome Reception</strong></td>
<td><strong>I-Hotel Chancellor Ballroom</strong></td>
</tr>
<tr>
<td>Monday, June 3</td>
<td><strong>Opening</strong></td>
<td><strong>Conference Welcome</strong></td>
</tr>
<tr>
<td>8:30 - 9:00am</td>
<td><strong>Keynote 1</strong></td>
<td><strong>Innovation is Dead! Long Live Innovation!</strong></td>
</tr>
<tr>
<td></td>
<td><strong>I-Hotel Chancellor Ballroom</strong></td>
<td><strong>Patricia Hswe</strong></td>
</tr>
<tr>
<td>9:00 - 10:30am</td>
<td><strong>Session 1</strong> 1A. Generation and Linking 1B. Analysis and Curation 1C. Search Logs</td>
<td><strong>I-Hotel Lincoln Room</strong></td>
</tr>
<tr>
<td></td>
<td><strong>I-Hotel Technology Room</strong></td>
<td><strong>I-Hotel Quad Room</strong></td>
</tr>
<tr>
<td></td>
<td><strong>I-Hotel Quad Room</strong></td>
<td></td>
</tr>
<tr>
<td>10:30 - 11:00am</td>
<td><strong>Break</strong></td>
<td></td>
</tr>
<tr>
<td>11:00am - 12:30pm</td>
<td><strong>Lunch</strong></td>
<td><strong>I-Hotel Chancellor Ballroom</strong></td>
</tr>
<tr>
<td>12:30 - 2:00pm</td>
<td><strong>Session 2</strong> 2A. Multimodal and Multimedia 2B. Classification, Discovery and Recommendation</td>
<td><strong>I-Hotel Lincoln Room</strong></td>
</tr>
<tr>
<td></td>
<td><strong>I-Hotel Technology Room</strong></td>
<td></td>
</tr>
<tr>
<td>2:00 - 3:30pm</td>
<td><strong>Break</strong></td>
<td></td>
</tr>
<tr>
<td>3:30 - 4:00pm</td>
<td><strong>Session 3</strong> 3A. Poster Madness</td>
<td><strong>I-Hotel Chancellor Ballroom</strong></td>
</tr>
<tr>
<td>4:00 - 5:00pm</td>
<td><strong>Poster Session and Reception</strong></td>
<td></td>
</tr>
<tr>
<td>5:00 - 7:00pm</td>
<td><strong>Tuesday, June 4</strong></td>
<td><strong>Conference Welcome</strong></td>
</tr>
<tr>
<td>8:30 - 9:00am</td>
<td><strong>Keynote 2</strong></td>
<td><strong>Standards and Communities: Connected People, Consistent Data, Usable Applications</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Rob Sanderson</strong></td>
<td><strong>I-Hotel Chancellor Ballroom</strong></td>
</tr>
<tr>
<td>9:00 - 10:30am</td>
<td><strong>Session 4</strong> 4A. Research and Open Data 4B. Web Archives 1 4C. Analysis and Processing</td>
<td><strong>I-Hotel Lincoln Room</strong></td>
</tr>
<tr>
<td></td>
<td><strong>I-Hotel Technology Room</strong></td>
<td><strong>I-Hotel Quad Room</strong></td>
</tr>
<tr>
<td>10:30 - 11:00am</td>
<td><strong>Break</strong></td>
<td></td>
</tr>
<tr>
<td>11:00am - 12:30pm</td>
<td><strong>Lunch</strong></td>
<td></td>
</tr>
<tr>
<td>12:30 - 2:00pm</td>
<td><strong>Lunch</strong></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Event</td>
</tr>
<tr>
<td>-------------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2:00 - 3:30</td>
<td>Session 5</td>
<td>5A. Scholary Documents</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5B. Web Archives 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5C. Panel: Institutionalizing and Sustaining Virtual Reality Experiences in Higher Education</td>
</tr>
<tr>
<td>3:30 - 4:00</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>4:00 - 5:00</td>
<td>Session 6</td>
<td>6A. Bibliographic Data and Matching</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6B. User Interface and Behavior</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6C. Citations</td>
</tr>
<tr>
<td>6:30 - 8:30</td>
<td>Banquet</td>
<td></td>
</tr>
</tbody>
</table>

**Wednesday, June 5**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 - 10:00</td>
<td>Keynote 3: How do you lift an elephant with one hand?</td>
<td>Chancellor Ballroom</td>
</tr>
<tr>
<td>10:00 - 10:30</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>10:30am - 12:00pm</td>
<td>7A. Panel: Towards a DL by the Communities and for the Communities</td>
<td>Chancellor Ballroom</td>
</tr>
<tr>
<td>12:00 - 12:30pm</td>
<td>Closing</td>
<td>Chancellor Ballroom</td>
</tr>
<tr>
<td>12:00 - 1:00pm</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>1:00 - 5:00pm</td>
<td>Workshop (1) Workshop on Requirements, Use Cases, and User Studies in Digital Music Libraries and Archives (RUCUS 2019)</td>
<td>I-Hotel Quad &amp; Technology Rooms</td>
</tr>
</tbody>
</table>

**Thursday, June 6**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00am - 12:00pm</td>
<td>Workshop (2): Web Archiving and Digital Libraries (WADL) 2019</td>
<td>Illini Union - Room 209</td>
</tr>
<tr>
<td></td>
<td>Workshop (3): Conceptual Models in Digital Libraries, Archives, and Museums</td>
<td>Illini Union - Room 404</td>
</tr>
<tr>
<td></td>
<td>Workshop (4): Organizing Data, Information, and Knowledge in Big Data Environments</td>
<td>Illini Union - Room 406</td>
</tr>
<tr>
<td></td>
<td>Tutorial (5): Text Mining with HathiTrust</td>
<td>Illini Union - Room 405</td>
</tr>
<tr>
<td>12:00 - 1:00pm</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>1:00 - 5:00pm</td>
<td>Workshop (2), cont. Web Archiving and Digital Libraries (WADL) 2019</td>
<td>Illini Union - Room 209</td>
</tr>
<tr>
<td></td>
<td>Workshop (3), cont. Conceptual Models in Digital Libraries, Archives, and Museums</td>
<td>Illini Union - Room 404</td>
</tr>
<tr>
<td></td>
<td>Workshop (4), cont. Organizing Data, Information, and Knowledge in Big Data Environments</td>
<td>Illini Union - Room 406</td>
</tr>
<tr>
<td></td>
<td>Tutorial (5), cont. Text Mining with HathiTrust</td>
<td>Illini Union - Room 405</td>
</tr>
</tbody>
</table>
Detailed Schedule

Sunday, June 2

8:00am-12:00pm
Tutorial (1)  
*Introduction to Digital Libraries*  
Edward A. Fox, Virginia Tech

*Doctoral Consortium*  
Oksana Zavalina, University of North Texas  
Peter Darch, University of Illinois at Urbana-Champaign  
Edie Rasmussen, University of British Columbia

12:00 - 1:00pm
Lunch Break  
*Lunch will be provided for Doctoral Consortium participants.*

1:00 - 5:00pm
Tutorial (2)  
*Strategies and Tools for Digital Repository Selection and Migration*  
Todd Crocken, University of Houston Libraries  
Anne Washington, University of Houston Libraries

Tutorial (3)  
*Building Digital Library Collections with Greenstone 3*  
David Bainbridge, University of Waikato, New Zealand

Tutorial (4)  
*Preparing Code and Data for Computational Reproducibility*  
April Clyburne-Sherin, Code Ocean  
Xu Fei, Code Ocean

*Doctoral Consortium, cont.*  
Oksana Zavalina, University of North Texas  
Peter Darch, University of Illinois at Urbana-Champaign  
Edie Rasmussen, University of British Columbia

5:00 - 6:00pm
Steering Committee Meeting

6:00 - 8:00pm
Welcome Reception
Monday, June 3

8:30 - 9:00am
Conference Welcome

9:00 - 10:30am
Keynote Speaker: Patricia Hswe
Innovation is Dead! Long Live Innovation!

10:30 - 11:00am
Break – Coffee and Tea

11:00am - 12:30pm
Session 1

1A. Generation and Linking
Automatic Generation of Initial Reading Lists: Requirements and Solutions
Pablo Figueira, Fabiano Belém, Jussara Almeida and Marcos Gonçalves

NAISC: An Authoritative Linked Data Interlinking Approach for the Library Domain
Lucy McKenna, Christophe Debruyne and Declan O'Sullivan

BioGen: Automated Biography Generation
Heer Ambavi, Ayush Garg, Ayush Garg, Nitiksha, Mridul Sharma, Rohit Sharma, Jayesh Choudhari and Mayank Singh

Corpus Assembly as Text Data Integration from Digital Libraries and the Web
Udo Hahn and Tinghui Duan

1B. Analysis and Curation
Deep Analysis of OCR Errors for Effective Post-OCR Processing
Thi Tuyet Hai Nguyen, Adam Jatowt, Mickaël Coustaty, Nhu Van Nguyen and Antoine Doucet

Digital Curation at Work: Modeling Workflows for Digital Archival Materials
Colin Post, Alexandra Chassanoff, Christopher Lee, Andrew Rabkin, Yinglong Zhang, Katherine Skinner and Sam Meister

Advancing Reproducibility Through Shared Data: Bridging Archival and Library Practice
Julianna Pakstis, Hannah Calkins, Christiana Dobrzynski, Spencer Lamm and Laura McNamara

Unsupervised Clustering with Smoothing for Detecting Paratext Boundaries in Scanned Documents
Ana Lucic, Robin Burke and John Shanahan
1C. Search Logs
Characterizing Searches for Mathematical Concepts
Behrooz Mansouri, Richard Zanibbi and Douglas Oard

Exploring Usage Patterns of a Large-Scale Digital Library
Maram Barifah and Monica Ladoni

Large Scale Analysis of Semantic and Temporal Aspects in Cultural Heritage Collection's Search
Yasunobu Sumikawa, Adam Jatowt, Antoine Doucet and Jean-Philippe Moreux

12:30 - 2:00pm
Lunch
Steering Committee Meeting

2:00 - 3:30pm
Session 2

2A. Multimodal and Multimedia
The Ownership and Control of Online Photos and Game Data: Patterns, Trends, and Keeping Pace with Evolving Circumstances
Catherine C. Marshall and Frank Shipman

Crawling and Classification Strategies for Generating a Multi-Language Corpus of Sign Language Video
Frank Shipman and Caio Monteiro

A Joint Model for Multimodal Document Quality Assessment
Aili Shen, Bahar Salehi, Timothy Baldwin and Jianzhong Qi

2B. Classification, Discovery and Recommendation
Learning from Few Samples: Lexical Substitution with Word Embeddings for Short Text Classification
Abel Elekes, Simone Di Stefano, Martin Schäler, Matthias Keller and Klemens Böhm

Improving Academic Plagiarism Detection for STEM Documents by Analyzing Mathematical Content and Citations
Norman Meuschke, Vincent Stange, Moritz Schubotz, Michael Kramer and Bela Gipp

Document Embeddings vs. Keyphrases vs. Terms for Recommender Systems: A Large-Scale Online Evaluation
Andrew Collins and Joeran Beel
‘Too Late to Collaborate’ Challenges to the Discovery of in-Progress Research
Corinna Breitinger, Patrick Wortner, Bela Gipp and Harald Reiterer

3:30 - 4:00pm
Break – Coffee and Tea

4:00 - 5:00pm
Session 3

3A. Poster Minute Madness
*See program for list of posters and demonstrations.

5:00 - 7:00pm
Poster Session and Reception

Tuesday, June 4

8:30 - 9:00am
Conference Announcements

9:00 - 10:30am
Keynote Speaker: Rob Sanderson
Standards and Communities: Connected People, Consistent Data, Usable Applications

10:30 - 11:00am
Break – Coffee and Tea

11:00am - 12:30pm
Session 4

4A. Research and Open Data
Modeling Digital Humanities Collections as Research Objects
Katrina Fenlon

A Digital Library for Research Data and Related Information in the Social Sciences
Daniel Hienert, Dagmar Kern, Katarina Boland, Benjamin Zapilko and Peter Mutschke

Open Data Publishing by Public Libraries
Nicholas Weber and Bree Norlander

4B. Web Archives 1
Archive Assisted Archival Ffixity Verification Framework
Mohamed Aturban, Sawood Alam, Michael Nelson and Michele Weigle
MementoMap Framework for Flexible and Adaptive Web Archive Profiling
Sawood Alam, Michele Weigle, Michael Nelson, Fernando Melo and Daniel Gomes

Evaluating Memento Service Optimizations
Martin Klein, Lyudmila Balakireva and Harihar Shankar

4C. Analysis and Processing
An End-to-End Approach for Extracting and Segmenting High-Variance References from PDF Documents
Zeyd Boukher, Shriharsh Ambhore and Steffen Staab

Automated Identification of Media Bias by Word Choice and Labeling in News Articles
Felix Hamborg, Anastasia Zhukova and Bela Gipp

Do Authors Deposit on Time? Tracking Open Access Policy Compliance
Drahomira Herrmannova, Nancy Pontika and Petr Knoth

12:30 - 2:00pm
Lunch

2:00 - 3:30pm
Session 5

5A. Scholarly Documents
What Drives Research Efforts? Find Scientific Claims that Count!
Jose Maria Gonzalez Pinto, Janus Wawrzeniek and Wolf-Tilo Balke

A Deep Multimodal Investigation To Determine the Appropriateness of a Scholarly Submission
Tirthankar Ghosal, Ashish Raj, Asif Ekbal, Sriparna Saha and Pushpak Bhattacharyya

Is the Paper Within Scope? Are You Fishing in the Right Pond?
Tirthankar Ghosal, Asif Ekbal, Sriparna Saha, Pushpak Bhattacharyya and Ravi Sonam

5B. Web Archives 2
Towards Temporal URI Collections for Named Entities
Sergej Wildemann and Helge Holzmann

Using Micro-Collections in Social Media to Generate Seeds for Web Archive Collections
Alexander Nwala, Michele Weigle and Michael Nelson

The Cost of a WARC: Analyzing Web Archives in the Cloud
Ryan Deschamps, Samantha Fritz, Jimmy Lin, Ian Milligan and Nick Ruest

Building Community and Tools for Analyzing Web Archives through Datathons
Ian Milligan, Nathalie Casemajor, Samantha Fritz, Jimmy Lin, Nick Ruest, Matthew S. Weber and Nicholas Worby
5C. Panel
Institutionalizing and Sustaining Virtual Reality Experiences in Higher Education
Jim Hahn and David Ward

3:30 - 4:00pm
Break – Coffee and Tea

4:00 - 5:00pm
Session 6

6A. Bibliographic Data and Matching
Dirichlet Process Gaussian Mixture for Active Online Name Disambiguation by Particle Filter
Baichuan Zhang, Murat Dundar, Vachik Dave and Mohammad Al Hasan
Author Matching Across Different Academic Databases: Aggregating Simple Feature-Based Rankings
Marie Katsurai and Ikki Ohmukai
Learning to Retrieve Related Resources in a Bibliographic Information Network
Poonam Anthony and Plaban Kumar Bhowmick

6B. User Interface and Behavior
One Way or Another I’m Gonna Find Ya: The Influence of Input Mechanism on Scrolling in Complex Digital Collections
George Buchanan and Dana Mckay
Recognizing Topic Change in Search Sessions of Digital Libraries Based on Thesaurus and Classification System
Daniel Hienert and Dagmar Kern
Shared Feelings: Understanding Facebook Reactions to Scholarly Articles
Curtis Cole Freeman, Mrinal Kanti Roy, Michele Fattoruso and Hamed Alhoori

6C. Citations
Go Wide, Go Deep: Quantifying the Impact of Scientific Papers through Influence Dispersion Trees
Dattatreya Mohapatra, Abhishek Maiti, Sumit Bhatia and Tanmoy Chakraborty
Venue Analytics: A Simple Alternative to Citation-Based Metrics
Leonid Keselman

6:30 - 8:30pm
Banquet
### Wednesday, June 5

<table>
<thead>
<tr>
<th>Time</th>
<th>Event Description</th>
</tr>
</thead>
</table>
| 8:30 - 10:00am | **Keynote Speaker:** John Wilkin  
*How do you lift an elephant with one hand?*  
**Chanellor Ballroom** |
| 10:00 – 10:30am | **Break – Coffee and Tea**                                                        |
| 10:30am - 12:00pm | **Session 7**  
**7A. Panel (Lincoln Room)**  
*Towards a DL by the Communities and for the Communities*  
Daqing He, Dan Wu, Wayne Graves, Martin Klein and Alexandra Dolan-Mescal |
| 12:00 - 1:00pm | **Lunch**  
**Chanellor Ballroom** |
| 1:00 – 5:00pm | **Workshop (1)**  
*Workshop on Requirements, Use Cases, and User Studies in Digital Music Libraries and Archives (RUCUS 2019)*  
David M. Weigl, University of Music and Performing Arts, Vienna, Austria  
**Quad & Technology Rooms** |

### Thursday, June 6

<table>
<thead>
<tr>
<th>Time</th>
<th>Event Description</th>
</tr>
</thead>
</table>
| 8:00am - 12:00pm | **Workshop (2)**  
*Web Archiving and Digital Libraries (WADL) 2019*  
Martin Klein, Los Alamos National Laboratory  
Zhiwu Xie, Virginia Tech  
Edward A. Fox, Virginia Tech  
**Illini Union Room 209** |
|              | **Workshop (3)**  
*Conceptual Models in Digital Libraries, Archives, and Museums*  
Nicholas Weber, University of Washington  
Katrina Fenlon, University of Maryland  
Peter Organisciak, University of Denver  
Andrea K. Thomer, University of Michigan  
**Illini Union Room 404** |
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
</table>
| 10:00 – 11:00 | Workshop (4), cont. *Organizing Data, Information, and Knowledge in Big Data Environments*  
Jiangping Chen, University of North Texas  
Wei Lu, Wuhan University  
Oksana L. Zavalina, University of North Texas | Illini Union  
Room 406 |
| 11:00 – 12:00 | Tutorial (5), cont. *Text Mining with HathiTrust*  
Eleanor Dickson Koehl, HathiTrust  
Ryan Dubnicek, University of Illinois at Urbana-Champaign | Illini Union  
Room 405 |
| 12:00 – 1:00  | Lunch                                                                 |                |
| 1:00 – 5:00   | Workshop (2), cont. *Web Archiving and Digital Libraries (WADL) 2019* | Illini Union  
Room 209 |
Room 404 |
|               | Workshop (4), cont. *Organizing Data, Information, and Knowledge in Big Data Environments* | Illini Union  
Room 406 |
|               | Tutorial (5), cont. *Text Mining with HathiTrust* | Illini Union  
Room 405 |
KEYNOTE SPEAKER
Monday, June 3

Innovation is Dead! Long Live Innovation!
Patricia Hswe
The Andrew W. Mellon Foundation
ph@mellon.org

ABSTRACT
It may be unthinkable to reconsider the concept and purpose of innovation and experimentation. In many sectors—including higher ed, corporations, the arts, humanities, and sciences—the creation of the Next New Shiny Thing can have its rewards, both financially and in terms of prestige. Experimental investigations are a key gateway to new knowledge; little learning and insight can happen without such developmental pursuits. At the same time, is there a way to recast innovation so that what is being built is not entirely new or has not been done before but, instead, makes adaptation, reuse, repair, recovery, and maintenance the overriding rationales for any “new” technology? If the needs of future users are at stake, particularly humanities scholars and students as users, then how should we in libraries, presses, archives, and information technology re-think innovation, so that what gets built and what has been built—what gets used and what has been used—evolve and endure rather than fade away and expire? How do we make maintenance the Next New Shiny Thing? This presentation will delve into issues of experimentation, innovation, continuity, and maintenance in the context of digital library (DL) infrastructure and technology, writ broadly, for the humanities. It will highlight DL efforts that present opportunities for recalibrating the meaning and ethos of “innovation,” and it will surface various current initiatives that may be considered models of maintenance.

BIOGRAPHY
Patricia Hswe (“sway”) is the program officer for Scholarly Communications at The Andrew W. Mellon Foundation, which she joined in August 2016. In this role she works on a range of grants and initiatives supporting libraries, archives, museums, universities, presses, and other institutions that further the world’s collective knowledge of the humanities. The areas of focus that Patricia is passionate about in grantmaking include community-based archiving; information and infrastructure maintenance for digital scholarship; collections-as-data approaches for use and reuse of cultural heritage content; and promoting diversity, equity, and inclusion principles in capacity building for collections, people, and systems. Previously, Patricia worked in the libraries at Penn State University, where she co-founded the department of publishing and curation services and at the University of Illinois at Urbana-Champaign, where she was project manager for grant initiatives funded by the Library of Congress and by the Institute of Museum and Library Services. Originally a Russian literature scholar, Patricia holds a PhD from Yale University in Slavic languages and literatures. She also received an AB in Russian language and literature from Mount Holyoke College and an MS in library and information science from the University of Illinois. Patricia is currently a member of the Executive Council of the Association for Computers and the Humanities.
KEYNOTE SPEAKER
Tuesday, June 4

Standards and Communities: Connected People, Consistent Data, Usable Applications
Robert Sanderson
J. Paul Getty Trust
RSanderson@getty.edu

ABSTRACT
Standards are one of the cornerstones of the connected world, allowing independently developed systems and data to interoperate without bilateral agreements. Yet standards are not worth the free GitHub repository they’re written in if no one implements them. What makes some specifications, endorsed by a recognized standards body or not, rise to the top and see broad adoption, while others languish, unloved, deep in the IETF, W3C, or ISO document management platforms? This presentation explores the connection between standards and the communities that need them, and the processes those communities and standards bodies have used to attempt to stay in touch with their potential user base. There is always a trade-off between completeness and usability, between production and consumption, between general market and specific needs of a community. The art of standards is walking the fine line for each of these, and having enough visibility in the community to have the right people at the table at the right time. With an emphasis on IIIF and Linked Open Data oriented technologies, but relating back to digital library efforts such as Z39.50, SRU and OAI-PMH, we will look at how the balance between community engagement, ease of implementation and usability needs to be balanced against meeting technical requirements and delivering a well-tested, broadly implemented ecosystem to enable researchers and enthusiasts to engage with cultural heritage and scientific literature.

BIOGRAPHY
Dr. Robert Sanderson is an internationally known information scientist and expert in Linked Open Data and cultural heritage standards. He is the J. Paul Getty Trust’s first Semantic Architect and is a passionate advocate for open digital cultural heritage. He is responsible for the design and direction of cultural heritage data information models and systems spanning the Museum, Research Institute, and Conservation Institute. His main goal is to find the right balance between ease of publication and consumption of data, and the precision of the data’s semantics. He is chair of the JSON-LD Working Group in the W3C, is a specification editor and leader in the IIIF community, and on the advisory boards of many projects in the cultural sector including the American Art Collaborative, Annotating All Knowledge, and Art Information Commons projects. He has previously been a Standards Advocate at Stanford University, a Research Scientist at Los Alamos National Laboratory’s Research Library, and a Lecturer in Computer Science at the University of Liverpool. He has been involved with numerous digital library standards and projects, and has over 150 publications and invited presentations to his name.
KEYNOTE SPEAKER  
Wednesday, June 5

How do you lift an elephant with one hand?  
John Wilkin  
University of Illinois at Urbana-Champaign  
jpwilkin@illinois.edu

ABSTRACT
There were tremendous challenges in creating HathiTrust. The resulting organization, digital preservation infrastructure, and collection contributes significant value to the library ecosystem. This shift to a complement of shared and local services argues for libraries to create more at-scale, shared initiatives to ensure a vital future for research libraries. We need to confront questions about why at-scale efforts are so uncommon.

BIO
John Wilkin assumed the position of Dean of Libraries and University Librarian at the University of Illinois at Urbana-Champaign in 2013. In addition, he served as Interim Vice Chancellor for Academic Affairs and Provost Designate at Illinois from February 2017 through January 2018. Previously, Wilkin served as Executive Director, HathiTrust and in a variety of administrative roles (including interim library director) at the University of Michigan. Under Wilkin’s leadership, Michigan’s Digital Library Production Service pioneered a number of large-scale digital library efforts, including Making of America, the Humanities Text Initiative, PEAK (a system delivering Elsevier’s journals), and putting the Middle English Dictionary online.
Posters and Demonstrations

Monday, June 3

4:00-5:00pm Poster Madness, I-Hotel Chancellor Ballroom
5:00-7:00pm Poster Session & Reception, I-Hotel Chancellor Ballroom

Posters

1) ACT: An Annotation Platform for Citation Typing at Scale
   David Pride, Jozef Harag and Petr Knoth

2) An Analysis of Arabic Language Music Queries: Design Considerations for an Arabic Music Digital Library
   Sally Jo Cunningham

3) An analysis of the performance of named entity recognition over OCRed documents
   Ahmed Hamdi, Axel Jean-Caurant, Nicolas Sidère, Mickaël Coustaty and Antoine Doucet

4) An Approach of Constructing Knowledge Graph of the Hundred Schools of Thought in Ancient China
   Jingzhu Wei and Rui Liu

5) The Archives Unleashed Notebook: Madlibs for Jumpstarting Scholarly Exploration
   Ryan Deschamps, Nick Ruest, Jimmy Lin, Samantha Fritz and Ian Milligan

6) Argument Structure Mining in Scientific Articles: A Comparative Analysis
   Ningyuan Song, Hanghang Cheng, Huimin Zhou and Xiaoguang Wang

7) Branch Filtering of Tree-structured Search Results
   Trond Aalberg

8) Building and Enhancing Stereographic Digital Collections Working Across Departments to Augment Reality
   Bryan Ricupero, Glory Taylor and Amanda Lehman

9) A Comparative Study of Chinese Patent Literature Automatic Classification Based on Deep Learning
   Lucheng Lyu and Tao Han

10) CrowdEIM: Crowdsourcing Emergency Information Management Tasks to the Mobile Social Media Users
    Hongzhou Shen and Junpeng Shi

11) A Dataset for Content Error Detection in Web Archives
    Johannes Kiesel, Fabienne Hubricht, Benno Stein and Martin Potthast

12) Developing Ontology for Digital Archives of Wartime Document Protection Projects
    Mei-Mei Wu, Ying-Hsang Liu and Pei-Hsin Chan

    Zhao Rongying, Li Xinlai and Li Danyang
14) DFS: A Dataset File System for Data Discovering Users
Yasith Jayawardana and Sampath Jayarathna

15) Differences in the Research Domains of Knowledge Organization between Academic Researchers and Library Practitioners: Preliminary Results
Soohyung Joo and Kyong Eun Oh

16) A Document Summary Method Based on the Relative Spatial Length between Sentences
Panote Siriaraya, Haruka Sakata, Yuanyuan Wang and Yukiko Kawai

17) Examining Patterns of Text Reuse in Digitized Text Collections
Peter Organisciak, Grace Therrell, Maggie Ryan and Benjamin Schmidt

18) Exploring the Impact of Behaviors on User’s Popularity in the SQA Community
Shengli Deng, Yuting Jiang and Sudi Xia

19) Exploring Trends in Open Access Repositories: The case of higher education institutions in Nigeria, Ghana, Cabo Verde, and Senegal
Zainab Iddriss and Amirah Al Sarraj

20) Extracting pest risk information from risk assessment documents
Glen Newton, Oksana Korol, Andre Levesque, Robert Favrin and Tom Graefenham

21) Global Science Discussed in Local Social Media: WeChat and its comparison with Web of Science
Ting Cong, Bikun Chen and Wei Ming

22) Global Science Discussed in Local Social Media: WeChat and its comparison with Web of Science
Ting Cong, Bikun Chen and Wei Ming

23) A graphical Interface for the Dodge Poetry Festival Archive
Nina Wacholder, Grey Valenti and Nicholas Provenzano

24) How to measure the consistency of the tagging of scientific papers?
Boris Veytsman

25) Incorporating Full Text and Bibliographic Features to Improve Scholarly Journal Recommendation
Tirthankar Ghosal, Ananya Chakraborty, Ravi Sonam, Asif Ekbal, Sriparna Saha and Pushpak Bhattacharyya

26) Information Seeking in Pokemon Go: A Preliminary Investigation
Krishna Srinivasan, Marcus Lee, Derrick Chin, Dion Goh and Chei Sian Lee

27) International Comparison of Digitized Book Index in Japan, the U.S.A, and the U.K.
Teru Agata and Shuichi Ueda

28) Less Than 10% of Library Service Users Ask For Help
Cristel Thomas and Martin Klein

29) Librarians’ Perceptions on Skills/Knowledge and Resources Needed for Research Data Services: Preliminary Results
Soohyung Joo and Christie Peters

30) Measuring the Interdisciplinary Degree of Information Behavior Research
Shengli Deng, Sudi Xia and Shaoxiong Fu
31) MinSciE: Citation-centered Open Information Extraction
   Anne Lauscher, Yide Song and Kiril Gashteovski

32) Mistaken Identity: Knowledge Service in China’s Academic Libraries
   Minglei Ying, Lihong Zhou and Ruhua Huang

33) Moves Recognition in Abstract of Research Paper Based on Deep Learning
   Zhixiong Zhang, Huan Liu and Liangping Ding

34) A Multiview Clustering Approach To Identify Out-of-Scope Submissions in Peer Review
   Tirthankar Ghosal, Debojyoti Dey, Avik Dutta, Asif Ekbal, Sriparna Saha and Pushpak Bhattacharyya

35) Neural Multi-Task Learning for Citation Function and Provenance
   Xuan Su, Animesh Prasad, Min-Yen Kan and Kazunari Sugiyama

36) Nucleus – Deploying Research Data Management Infrastructure At The Los Alamos National Laboratory
   Brian Cain, Martin Klein, and Joshua Finnell

37) Practice of Constructing Name Authority Database Based on Multi-Source Data Integration
   Qianqian Yu, Jianyong Zhang, Li Qian, Zhipeng Dong, Yongwen Huang and Jianhua Liu

38) Predicting Patent Citations to measure Economic Impact of Scholarly Research
   Abdul Rahman Shaikh and Hamed Alhoori

39) A Ranking Method for Relaxed Queries in Book Search
   Momo Kyozuka, Yang Xu and Keishi Tajima

40) Recommending Personalized Search Terms for Assisting Exploratory Website Search
   Young Park

41) ReTracker: an Open-source Plugin for Automated and Standardized Tracking of Retracted Scholarly Publications
   Ly Dinh, Yi-Yun Cheng and Nikolaus Parulian

42) The Role of Self-Efficacy and Familiarity in Digital Humanity Crowdsourcing: A Preliminary Study from Transcribe-Sheng Project
   Wenting Han, Shijie Song, Yuxiang Zhao and Qinghua Zhu

43) Semantically Enriched Presentation for Cultural Heritage Image: A POI-Based Perspective
   Xiaoxi Luo, Xu Tan and Xiaoguang Wang

44) Semantifying the UK Hansard (1918-2018)
   Federico Nanni, Stefano Menini, Sara Tonelli and Simone Paolo Ponzetto

45) A Sentiment Augmented Deep Architecture to Predict Peer Review Outcomes
   Tirthankar Ghosal, Rajeev Verma, Asif Ekbal and Pushpak Bhattacharyya

46) To Help without Expectation: Investigating Social Exchanges on a Mobile Crowdsourcing Platform for a Smart City
   Chei Sian Lee, Dion Goh and Hamzah Osop

47) Topic Detection of Online Book Reviews: Preliminary Results
   Yunseon Choi and Soohyung Joo

48) Transaction Log Analysis within a Bento Discovery System
   Willaim Mischo, Michael Norman and Mary Schlembach
49) **Using the HTRC Data Capsule Model to Promote Reuse and Evolution of Experimental Analysis of Digital Library Data: A Case Study of Topic Modeling**
   David Bainbridge, David Nichols, Annika Hinze and J. Stephen Downie

50) **Using VIAF Dataset and the National Bibliography for Identifying and Listing Comics and Manga Authors**
   Yasuharu Otani, Teru Agata, Akiko Hashizume, Masaki Eto, Mari Agata and Noriko Sugie

51) **Web Archive Analysis Using Hive and SparkSQL**
   Xinyue Wang and Zhiwu Xie

52) **When Should FAIR Begin? Applying FAIR During Data Creation**
   Michael Gryk, Adam Schuyler, Jon Wedell, Mark Maciejewski, Gerard Weatherby and Pedro Romero

**Accepted Demonstrations**

53) **ChairVisE: An analytic lens for conference submission data**
   Animesh Prasad, Saumya Ahuja, Shenhao Jiang, Bimlesh Wadhwa and Min-Yen Kan

54) **EAL: A Toolkit and Dataset for Entity-Aspect Linking**
   Federico Nanni, Jingyi Zhang, Ferdinand Betz and Kiril Gashteovski

55) **EXCITE – A toolchain to extract, match and publish open literature references**
   Azam Hosseini, Behnam Ghavimi and Philipp Mayr

   Kevin Page, David Lewis and David M. Weigl

57) **Scalable Content-Based Analysis of Images in Web Archives with TensorFlow and the Archives Unleashed Toolkit**
   Hsiu-Wei Yang, Linqing Liu, Ian Milligan, Nick Ruest and Jimmy Lin

58) **ScholarSight: Visualizing Temporal Trends of Scientific Concepts**
   Michael Färber, Chifumi Nishioka and Adam Jatowt

59) **Tempurion: A Collaborative Temporal URI Collection for Named Entities**
   Sergej Wildemann and Helge Holzmann

60) **Warclight: A Rails Engine for Web Archive Discovery**
   Nick Ruest, Ian Milligan and Jimmy Lin
Sponsors

Diamond Supporter:

Platinum Supporter:
Gold Supporters:

University of Illinois System

School of Information Management Wuhan University

University of Virginia Library

Elsevier

Silver Supporters:

Department of Information Science
College of Information
UNT EST. 1890

University of Pittsburgh School of Computing and Information

HathiTrust Research Center
Bronze Supporters:

- ILLINOIS
  NCSA | National Center for Supercomputing Applications

- ILLINOIS
  Technology Services | RESEARCH IT

- IEEE Xplore®
  Digital Library

- UNC | SCHOOL OF INFORMATION AND LIBRARY SCIENCE

- THE UNIVERSITY OF TENNESSEE
  KNOXVILLE
  SCHOOL OF INFORMATION SCIENCES

- University Libraries
  UNIVERSITY OF COLORADO Boulder
Welcome to Champaign County!

Champaign County offers an Outside of Ordinary experience, with the amenities of a big city with the charm of a small town. Discover the many bars and restaurants, shops, galleries, theaters, attractions and more in our walkable districts, as well as all of our hotels and attractions county-wide.

For more information on all there is to experience in Champaign County: visitchampaigncounty.org :: 800.369.6151

For more information on promoting your business with the Champaign Center Partnership: champaigncenter.com :: 217.352.2400

Visit Champaign County

Map Key

Denotes Champaign Center Partnership Members

Champaign Center Partnership is a local nonprofit business association serving Champaign’s Downtown, Midtown and Campustown neighborhoods. Learn more about the benefits of joining their growing community of 130+ businesses and organizations at ChampaignCenter.com.

Disclaimer: This map was created in a joint partnership with Visit Champaign County and the Champaign Center Partnership (CCP). The map includes many shops, bars, restaurants, theaters and galleries in Downtown, Midtown and Campustown neighborhoods. Learn more about the benefits of joining their growing community of 130+ businesses and organizations at ChampaignCenter.com.

Food & Drink

1. 303 Kitchen
2. 303 Bistro
3. 401 Brew Lab Coffee
4. 505 Empanadas House
5. 606 Dunkin’ Donuts
6. 707 Caffe Bene
7. 808 Fat Sandwich Company
8. 909 Fat City Bar & Grill

Accommodations

1. 101 Bellmore Indian Grill
2. 102 Howard Johnson
3. 103 Holiday Inn
4. 104 The Hub Sports Bar and Grill
5. 105 Potbelly’s Subs & Sandwiches
6. 106 Red Lion
7. 107 Stoney’s Bar & Grill
8. 108 The Spot

Shopping

1. 109 Apricot Lane
2. 110 The Illini Shop
3. 111 In Stock
4. 112 The Illini Shop

Theaters, Art & Museums

1. 113 Champaign Central Partnership: champaigncenter.com :: 217.352.2400
2. 114 CAMPUSTOWN
3. 115 University Avenue

Wellness & Personal Care

1. 116 CAMPUSTOWN
2. 117 University Avenue

Other

1. 118 CAMPUSTOWN
2. 119 University Avenue

Financial Institutions

1. 120 CAMPUSTOWN
2. 121 University Avenue

Credit Union

1. 122 CAMPUSTOWN
2. 123 University Avenue

Accommodations

1. 124 CAMPUSTOWN
2. 125 University Avenue

Food & Drink

1. 126 CAMPUSTOWN
2. 127 University Avenue

Shopping

1. 128 CAMPUSTOWN
2. 129 University Avenue

Theaters, Art & Museums

1. 130 CAMPUSTOWN
2. 131 University Avenue

Wellness & Personal Care

1. 132 CAMPUSTOWN
2. 133 University Avenue

Other

1. 134 CAMPUSTOWN
2. 135 University Avenue

Financial Institutions

1. 136 CAMPUSTOWN
2. 137 University Avenue

The Champaign Center Partnership is a local nonprofit business association serving Champaign’s Downtown, Midtown and Campustown neighborhoods. Learn more about the benefits of joining their growing community of 130+ businesses and organizations at ChampaignCenter.com.

DISCLAIMER: This map was created in a joint partnership with Visit Champaign County and the Champaign Center Partnership (CCP). The map includes many shops, bars, restaurants, theaters and galleries in Downtown, Midtown and Campustown neighborhoods. Learn more about the benefits of joining their growing community of 130+ businesses and organizations at ChampaignCenter.com.

The Champaign Center Partnership is a local nonprofit business association serving Champaign’s Downtown, Midtown and Campustown neighborhoods. Learn more about the benefits of joining their growing community of 130+ businesses and organizations at ChampaignCenter.com.

See reverse side for Downtown Champaign and an area map of Champaign County.

Visit Champaign County :: 800.369.6151 /visitchampaign :: #outsideofordinary

Champaign County

Visit Champaign County

Champaign County

Visit Champaign County

Champaign County

Visit Champaign County

Champaign County