

2014 IEEE/ACM Joint Conference on Digital Libraries (JCDL 2014)

**London, United Kingdom
8-12 September 2014**



**IEEE Catalog Number: CFP14JCD-POD
ISBN: 978-1-4799-5570-1**

2014 IEEE/ACM Joint Conference on Digital Libraries (JCDL)

Table of Contents

	Page
Section: Preservation strategies	
When Should I Make Preservation Copies of Myself? <i>Chuck Cartledge, Michael Nelson</i>	1
An Argument for Archiving Facebook as a Heterogeneous Personal Store <i>Catherine Marshall, Frank Shipman</i>	11
Implementing Digital Preservation Strategy: Developing Content Collection Profiles at the British Library <i>Michael Day, Ann MacDonald, Akiko Kimura, Maureen Pennock</i>	21
The Archival Acid Test: Evaluating Archive Performance on Advanced HTML and JavaScript <i>Mat Kelly, Michael L. Nelson, Michele C. Weigle</i>	25
Section: Recommendation and indexing	
Recommendation Based on Deduced Social Networks in an Educational Digital Library <i>Monika Akbar, Clifford Shaffer, Weiguo Fan, Edward Fox</i>	29
Dynamic Taxonomy Composition via Keyqueries <i>Tim Gollub, Matthias Hagen, Michael Völske, Benno Stein</i>	39
Personalized PageRank for Making Recommendations in Digital Cultural Heritage Collections <i>Arantxa Otegi, Eneko Agirre, Paul Clough</i>	46
Making Research Data Findable in Digital Libraries: A Layered Model for User-Oriented Indexing of Survey Data <i>Tanja Friedrich, Andreas Oskar Kempf</i>	53
Session: Publication impacts	
Characterizing Scholar Popularity: A Case Study in the Computer Science Research Community <i>Glauber Dias Gonçalves, Flavio Vinicius Diniz de Figueiredo, Jussara Marques de Almeida, Marcos Andre Goncalves</i>	57
Community-based Endogamy as an Influence Indicator <i>Thiago H. P. Silva, Mirella M. Moro, Ana Paula C. Silva, Wagner Meira Jr., Alberto Laender</i>	67
Disambiguating Publication Venue Titles using Association Rules <i>Denilson Pereira, Eduardo Silva, Ahmed Esmin</i>	77
Session: Personal DL Design	
From User Needs to Opportunities in Personal Information Management: A Case Study on Organisational Strategies in Cross-Media Information Spaces <i>Sandra Trullemans, Beat Signer</i>	87
PerCon: A Personal Digital Library for Heterogeneous Data <i>Su Inn Park, Frank Shipman</i>	97
Social Information Behaviour in Physical Libraries: Implications for the design of digital libraries <i>Annika Hinze, Claire Timpany, Nicholas Vanderschantz, Hayat Alqurashi</i>	107

	Page
Session: Building systems	
Towards Building a Scholarly Big Data Platform: Challenges, Lessons and Opportunities <i>Zhaohui Wu, Jian Wu, Madian Khabza, Kyle Williams, Hung-Hsuan Chen, Wenyi Huang, Suppawong Tuarob, Sagnik Ray Choudhury, Alexander Ororbia, Prasenjit Mitra, C. Lee Giles</i>	117
Bridging the Gap Between Real World Repositories and Scalable Preservation Environments <i>Bolette Ammitzbøll Jurik, Asger Askov Blekinge, Rune Bruun Ferneke-Nielsen, Per Møldrup-Dalum</i>	127
Using ACM DL paper metadata as an auxiliary source for building educational collections <i>Yinlin Chen, Edward Fox</i>	137
Crowd-sourcing Web Knowledge for Metadata Extraction, <i>Zhaohui Wu, Wenyi Huang, Chen Liang, C. Lee Giles</i>	141
Session: Browsing and searching	
Lend me some sugar: Borrowing rates of neighbouring books as evidence for browsing <i>Dana McKay, Wally Smith, Shanton Chang</i>	145
Improving the visibility of geospatial data on the Web <i>Javier Lacasta, Francisco Javier López-Pellicer, Walter Renteria-Agualimpia, Javier Nogueras-Iso</i>	155
The Anatomy of a Search and Mining System for Digital Humanities <i>Martyn Harris, Mark Levene, Dell Zhang, Dan Levene</i>	165
Increasing the visibility of library records via a consortial search engine <i>Önne Mets, Silvia Gstrein, Veronika Gründhammer</i>	169
Session: Item identification	
Combining Domain-Specific Heuristics for Author Name Disambiguation <i>Alan Filipe Santana, Marcos André Gonçalves, Anderson A. Ferreira, Alberto Laender</i>	173
Detecting and Modeling Local Text Reuse <i>David A. Smith, Ryan Cordell, Elizabeth Maddock Dillon, John Wilkerson, Nick Stramp</i>	183
Identifying the Same Records across multiple Ukiyo-e Image Databases Using Textual Data in Different Languages <i>Biligsaikhan Batjargal, Takeo Kuyama, Fuminori Kimura, Akira Maeda</i>	193
Reducing Computational Effort for Plagiarism Detection by using Citation Characteristics to Limit Retrieval Space <i>Norman Meuschke, Bela Gipp</i>	197
Session: Quality Data and Metadata	
Quality Assessment of Collaborative Content With Minimal Information <i>Daniel Hasan Dalip, Harley Lima, Marcos Gonçalves, Marco Cristo, Pável Calado</i>	201
Fast Image-based Chinese Calligraphic Character Retrieval on Large Scale Data <i>Pengcheng Gao, Jiangqin Wu, Yuan Lin, Yang Xia, Tianjiao Mao, Baogang Wei</i>	211
Human and Machine Error Analysis on Dependency Parsing of Ancient Greek Texts <i>Saeed Majidi, Gregory Crane</i>	221
The Feasibility of Investing of Manual Correction of Metadata for a Large-Scale Digital Library	225

	Page
Session: Topics, evolution, and relationships	
A Framework for Analyzing Semantic Change of Words across Time <i>Adam Jatowt, Kevin Duh</i>	229
Representing Topics Labels for Exploring Digital Libraries <i>Nikolaos Aletras, Timothy Baldwin, Jey Han Lau, Mark Stevenson</i>	239
Topical Establishment Leveraging Literature Evolution <i>Han Xu, Eric Martin, Ashesh Mahidadia</i>	249
A Method to Support Analysis of Personal Relationship through Place Names Extracted from Documents <i>Fuminori Kimura, Akira Maeda</i>	253
Session: Knowledge Infrastructure and Repositories	
The Ups and Downs of Knowledge Infrastructures in Science: Implications for Data Management <i>Christine L. Borgman, Peter T. Darch, Ashley E. Sands, Jillian C. Wallis, Sharon Traweek</i>	257
CED2AR: The Comprehensive Extensible Data Documentation and Access Repository <i>Carl Lagoze, Lars Vilhuber, Jeremy Williams, Benjamin Perry, William C. Block</i>	267
Big Brother is Watching You -- But in a Good Way <i>Carllin St Pierre, David Bainbridge, Bill Rogers</i>	277
A comparative analysis of the HSS & HEP data submission workflows <i>Suenje Dallmeier Tiessen, Artemis Lavasa, Laura Rueda, Patricia Heterich, Rachael Kotarski, Elizabeth Newbold</i>	281
An Open Cultural Digital Content Infrastructure <i>Ioanna Ourania Stathopoulou, Haris Georgiadis, Vangelis Banos, Panagiotis Stathopoulos, Nikos Houssos, Evi Sachini</i>	285
Session: Data transformation and description	
Converting the Zeri photo archive in Linked Open Data: formalizing the conceptual model <i>Ciro Gonano, Francesca Mambelli, Silvio Peroni, Francesca Tomasi, Fabio Vitali</i>	289
Comic2CEBX: A System for Automatic Comic Content Adaptation <i>Luyuan Li, Yongtao Wang, Liangcai Gao, Zhi Tang, Ching Y. Suen</i>	299
Explorations in Linked Data practice for early music corpora <i>Tim Crawford, Ben Fields, David Lewis, Kevin R. Page</i>	309
Creating lightweight ontologies for dataset description: Practical applications in a cross-domain research data management workflow <i>João Aguiar Castro, João Rocha Da Silva, Cristina Ribeiro</i>	313
A preliminary evaluation of HathiTrust metadata: Assessing the sufficiency of legacy records <i>Katrina Fenlon, Colleen Fallaw, Timothy W. Cole, Myung-Ja Han</i>	317
Session: Web archives and memory	
Not All Mementos Are Created Equal: Measuring The Impact Of Missing Resources <i>Justin F. Brunelle, Mat Kelly, Hany Salaheldeen, Michele C. Weigle, Michael L. Nelson</i>	321
Finding Pages on the Unarchived Web <i>Hugo Huurdeman, Anat Ben-David, Jaap Kamps, Thaer Samar, Arjen P. de Vries</i>	331
What Triggers Human Remembering of Events? A Large-Scale Analysis of Catalysts for Collective Memory in Wikipedia	341

	Page
Session: Citation, citation, citation	
Towards a Stratified Learning Approach to Predict Future Citation Counts <i>Tanmoy Chakraborty, Suhansanu Kumar, Pawan Goyal, Niloy Ganguly, Animesh Mukherjee</i>	351
Full-Text based Context-Rich Heterogeneous Network Mining Approach for Citation Recommendation <i>Xiaozhong Liu, Yingying Yu, Chun Guo, Yizhou Sun, Liangcai Gao</i>	361
RefSeer: A Citation Recommendation System <i>Wenyi Huang, Zhaohui Wu, Prasenjit Mitra, C. Lee Giles</i>	371
Do Altmetrics Follow the Crowd or Does the Crowd Follow Altmetrics? <i>Hamed Alhoori, Richard Furuta</i>	375
Session: Education and collaboration	
Towards Automatic Identification of Core Concepts in Educational Resources <i>Md Sultan, Steven Bethard, Tamara Sumner</i>	379
Using Affective Embodied Agents in Information Literacy Education <i>Yanru Guo, Dion Hoe-Lian Goh, Brendan Luyt</i>	389
Bend Me Shape Me: A Practical Experience of Repurposing Research Data <i>Dana McKay</i>	399
Research networks in data repositories <i>Mark R. Costa, Jian Qin, Jun Wang</i>	403
Demonstrations	
TagTick: A Tool for Annotation Tagging over Solr indexes <i>Michele Artini, Claudio Atzori, Alessia Bardi, Sandro La Bruzzo, Paolo Manghi</i>	407
Keeping your Aggregative Infrastructure Under Control <i>Michele Artini, Claudio Atzori, Paolo Manghi</i>	409
Epimenides: An Information System offering Automated Reasoning for the Needs of Digital Preservation <i>Yannis Kargakis, Yannis Tzitzikas</i>	411
Extraction of Evolution Descriptions from the Web <i>Helge Holzmann, Thomas Risse</i>	413
Rexplore: Unveiling the Dynamics of Scholarly Data <i>Francesco Osborne, Enrico Motta</i>	415
Explore The Stacks: A System for Exploration in Digital Libraries <i>Mark Michael Hall</i>	417
Great War Stories Told by the People - Crowdsourced Cultural Heritage in Digital Museums <i>Ingo Frommholz, David Graves, Haiming Liu, Ashwin Kumar, Gordon Brady</i>	419
When Catalogs Collide: A Mashup Up of the Bibliographic Records from New Zealand's National Bibliography and the HathiTrust <i>Steffan Safey, David Bainbridge</i>	421
LODE: Linking and Enhancing High Quality RDF Repositories to the Web of Data <i>Timo Szttyler, Jakob Huber, Jan Noessner, Jaimie Murdock, Colin Allen, Mathias Niepert</i>	423
The SCAPE preservation lifecycle <i>Kresimir Duretec, Michael Kraxner, Artur Kulmukhametov, Markus Plangg, Christoph Becker, Luis Faria</i>	425
SNAC: The Social Networks and Archival Context Project - Towards an Archival Authority Cooperative <i>Ray Larson, Daniel Pitti, Adrian Turner</i>	427

	Page
Demonstrations (continued)	
Visualized Related Topics (VRT) System for Health Information Retrieval <i>Sukjin You, Joel Desarmo, Xiangming Mu, Sukwon Lee, Jessica C. Neal</i>	429
Posters	
Modeling Abstractions for Dance Digital Libraries <i>Katerina El Raheb, Yannis Ioannidis</i>	431
Text touching effects: Why it is difficult to do active reading with a touch-based tablet device <i>Hirohito Shibata, Kentaro Takano</i>	433
Vector-Borne Disease Network Digital Library <i>Michelle Barker, Donald Brower, Natalie Meyers</i>	435
CKGHV: a Comprehensive Knowledge Graph For History Visualization <i>Yingzhen Zhu, Xinyi Cao, Yali Bian, Jiangqin Wu</i>	437
The Value of Risk Management for Data Management in Science and Engineering <i>Filipe Ferreira, Ricardo Vieira, José Borbinha</i>	439
Quantifying the State Trials <i>Tracy Bergstrom, Donald Brower, Natalie Meyers, Sonia Howell, Eric Morgan, Elliott Visconsi</i>	441
On Cloud Deployment of Digital Preservation Environments <i>Daniel Pop, Marian Neagul, Dana Petcu</i>	443
Microscopic analysis of document handling while reading paper documents to improve digital reading device <i>Kentaro Takano, Hirohito Shibata, Junko Ichino, Tomonori Hashiyama, Shun'ichi Tano</i>	445
Amplifying Scientific Paper's Abstract By Leveraging Data-weighted Reconstruction <i>Shansong Yang, Weiming Lu, Baogang Wei, Wenjia An</i>	447
Unified Digital Library and Data Access Framework For Heterogeneous and Federated Computational Epidemiology Datasets <i>S.M.Shamimul Hasan, Sandeep Gupta, Edward A. Fox, Keith Bisset, Madhav V. Marathe</i>	449
Large-Scale Extraction and Analysis of Web Link References in Scholarly Articles <i>Ke Zhou, Richard Tobin, Claire Grover</i>	451
What Is This Song about Anyway?: Automatic Classification of Aboutness Using User Interpretations, Social Tags, and Lyrics <i>Kahyun Choi, Jin Ha Lee, J. Stephen Downie</i>	453
REEL: A Relation Extraction Learning Framework <i>Pablo Barrio, Gonçalo Simões, Helena Galhardas, Luis Gravano</i>	455
Linking Thesaurus for the Social Sciences to the Web of Linked Data <i>Andias Wira Alam, Andreas Oskar Kempf, Benjamin Zapilko</i>	457
A Context Model for Digital Preservation of Processes and its Application to a Digital Library System <i>Rudolf Mayer, Andreas Rauber, Gonçalo Antunes</i>	459
The Organization information integration in the management of a Digital Library System <i>Angela Di Iorio, Marco Schaerf</i>	461
Life span of web pages: A survey of 10 million pages collected in 2001 <i>Teru Agata, Yosuke Miyata, Emi Ishida, Atsushi Ikeuchi, Shuichi Ueda</i>	463
PageRank-based Word Sense Induction within Web Search Results Clustering <i>Jose Moreno, Gaël Dias</i>	465

	Page
Posters (Continued)	
Enabling Multilingual Information Access to Digital Collections: An Investigation of Metadata Records Translation <i>Jiangping Chen, Olajumoke Azogu, Ryan Knudson</i>	467
Mink: Integrating the Live and Archived Web Viewing Experience Using Web Browsers and Memento <i>Mat Kelly, Michael L. Nelson, Michele C. Weigle</i>	469
Cross-Cultural Mood Regression for Music Digital Libraries <i>Xiao Hu, Yi-Hsuan Yang</i>	471
Establishing an Online Access Panel for Interactive Information Retrieval Research <i>Dagmar Kern, Peter Mutschke, Philipp Mayr</i>	473
Mood Metadata for Video Games and Interactive Media <i>Stephanie Rossi, Jin Ha Lee, Rachel Ivy Clarke</i>	475
Balancing Factors Affecting Virtual Reference Services: Identified from Academic Librarians' Perspective <i>Sukjin You, Joel Desarmo, Xiangming Mu, Alexandra Dimitroff</i>	477
Articles, Papers, Chapters, Theses – who wins the Visibility Wars? <i>Melius Weideman.</i>	479
Exploring Relationships Among Video Games <i>Rachel Ivy Clarke, Jin Ha Lee, Jacob Jett, Simone Sacchi</i>	481
Where should I publish?--Detecting journal similarity based on what have been published there <i>Rob Koopman, Shenghui Wang</i>	483
Organization Structures of Personal Information: A Comparison on Two Groups of Information Workers <i>Hong Zhang, Xiao Hu</i>	485
A computational approach to understanding and predicting the behavior of educators using an online curriculum planning tool <i>Ogheneovo Dibia, Keith Maull, Tamara Sumner</i>	487
An Approach to Named Entity Extraction from Historical Documents in Traditional Mongolian Script <i>Biligsaikhan Batjargal, Garmaabazar Khaltarkhuu, Fuminori Kimura, Akira Maeda</i>	489
Europeana Cloud: Towards a Shared Cloud Computing Infrastructure for European Aggregators <i>Pavel Kats, Petr Knoth, Georgios Mamakis, Marcin Mielnicki, Markus Muhr, Marcin Werla</i>	491
A Shared Language for Building a Dataset of Sensitive Information <i>Clare Llewellyn, Laine Ruus, Ros Burnett, Steve Kirkwood, Mark Smith, Rocio Von-Jungenfeld</i>	493