2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL 2021)

Virtual Conference 27-30 September 2021



IEEE Catalog Number: CFP21JCD-POD ISBN: 978-1-6654-1771-6

Copyright © 2021 by the Institute of Electrical and Electronics Engineers, Inc. All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

*** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.

 IEEE Catalog Number:
 CFP21JCD-POD

 ISBN (Print-On-Demand):
 978-1-6654-1771-6

 ISBN (Online):
 978-1-6654-1770-9

Additional Copies of This Publication Are Available From:

Curran Associates, Inc 57 Morehouse Lane Red Hook, NY 12571 USA Phone: (845) 758-0400

Fax: (845) 758-2633

E-mail: curran@proceedings.com Web: www.proceedings.com



2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL) JCDL 2021

Table of Contents

Message from the General Chairxiv
Message from the Program Committee Chairsxv
Organizing Committee xvi
Program Committee xvii
Doctoral Consortiumxxi
Keynotes xxiii
Sponsorsxxv
JCDL 2021 Full Papers
A Comparative Analysis of Article Recommendation Platforms
A New Methodology to Bring Out Typical Users Interactions in Digital Libraries
A Toolbox for the Nearly-Unsupervised Construction of Digital Library Knowledge Graphs 21 Hermann Kroll (TU Braunschweig, Germany), Jan Pirklbauer (TU Braunschweig, Germany), and Wolf-Tilo Balke (TU Braunschweig, Germany)
Automatic Metadata Generation for Fish Specimen Image Collections
Comparing Personalized PageRank and Activation Spreading in Wikipedia Diagram-Based Search. 41 Hisham Benotman (Purdue University) and David Maier (Portland State University)

Diachronic Analysis of German Parliamentary Proceedings: Ideological Shifts through the Lens of Political Biases	51
Tobias Walter (University of Mannheim, Germany), Celina Kirschner (University of Mannheim, Germany), Steffen Eger (Technische	
Universität Darmstadt, Germany), Goran Glavaš (University of Mannheim, Germany), Anne Lauscher (University of Mannheim, Germany), and Simone Paolo Ponzetto (University of Mannheim, Germany)	
Do You Think It's Biased? How To Ask For The Perception Of Media Bias	61
Estimating Contemporary Relevance of Past News	70
Garbage, Glitter, or Gold: Assigning Multi-Dimensional Quality Scores to Social Media Geeds for Web Archive Collections Alexander C. Nwala (Indiana University, USA), Michele C. Weigle (Old Dominion University, USA), and Michael L. Nelson (Old Dominion University, USA)	80
GraphConfRec: A Graph Neural Network-Based Conference Recommender System	90
mproved Discoverability of Digital Objects in Institutional Repositories Using Controlled Vocabularies	00
Bertha Chipangila (University of Zambia, Zambia), Erick Liswaniso (University of Zambia, Zambia), Andrew Mawila (University of Zambia, Zambia, Zambia), Philomena Mwanza (University of Zambia, Zambia), Daisy Nawila (University of Zambia, Zambia), Robert M'sendo (University of Zambia, Zambia, Zambia), Mayumbo Nyirenda (University of Zambia, Zambia), and Lighton Phiri (University of Zambia, Zambia)	
t's All About The Cards: Sharing on Social Media Encouraged HTML Metadata Growth	10
Museum Experience into a Souvenir: Generating Memorable Postcards from Guide Device Behavior Log	20

Newsalyze: Effective Communication of Person-Targeting Biases in News Articles	. 130
Germany)	
NoteLink: A Point-and-Shoot Linking Interface between Students' Handwritten Notebooks and Instructional Videos	. 140
Profiling Web Archival Voids for Memento Routing	. 150
Replaying Archived Twitter: When your Bird is Broken, will it Bring you down?	160
S2AND: A Benchmark and Evaluation System for Author Name Disambiguation	. 170
ScanBank: A Benchmark Dataset for Figure Extraction from Scanned Electronic Theses and	
Dissertations	180
Scientific Data Management for Interconnected Critical Infrastructure Systems	192
Surfacing Collective Harms in Privacy Sensitive Data	202
Visualizing Feature-Based Similarity for Research Paper Recommendation	. 212
JCDL 2021 Short Papers	
A Deep Neural Architecture for Decision-Aware Meta-Review Generation	222

Are Neural Language Models Good Plagiarists? A Benchmark for Neural Paraphrase Detection Jan Philip Wahle (University of Wuppertal, Germany), Terry Ruas (University of Wuppertal, Germany), Norman Meuschke (University of Wuppertal, Germany), and Bela Gipp (University of Wuppertal, Germany)	. 226
Automatic Metadata Extraction Incorporating Visual Features from Scanned Electronic Theses and Dissertations Muntabir Hasan Choudhury (Old Dominion University, USA), Himarsha R. Jayanetti (Old Dominion University, USA), William A. Ingram (Virginia Polytechnic Institute and State University, USA), Jian Wu (Old Dominion University, USA), and Edward A. Fox (Virginia Polytechnic Institute and State University, USA)	. 230
Change Summarization of Diachronic Scholarly Paper Collections by Semantic Evolution Analysis Naman Paharia (Indian Institute of Technology, India), Muhammad Syafiq Mohd Pozi (Universiti Utara, Malaysia), and Adam Jatowt (University of Innsbruck, Austria)	. 234
COMPARE: A Taxonomy and Dataset of Comparison Discussions in Peer Reviews	238
DRESS: Data-Repository Enhancer Through Semantic Sources Angel L. Garrido (University of Zaragoza, Spain) and Carlos Bobed (University of Zaragoza, Spain)	. 242
Exploring the Classification of Traditional Chinese Bibliographies Through Interactive Visualization Wenqi Li (Peking University, China), Fengxiang Wang (Peking University, China), and Jun Wang (Peking University, China)	. 246
MexPub: Deep Transfer Learning for Metadata Extraction from German Publications	250
Pathways to Data: From Plans to Datasets	. 254
Recognize, Annotate, and Visualize Parallel Content Structures in XML Documents Marco Beck (University of Wuppertal, Germany), Moritz Schubotz (FIZ Karlsruhe, Germany), Vincent Stange (OriginStamp AG, Switzerland), Norman Meuschke (University of Wuppertal, Germany), and Bela Gipp (University of Wuppertal, Germany)	258
References of References: How Far is the Knowledge Ancestry	. 262

Resource Types Linked in Academic Reading Lists Nandana Kumara (University of Waikato, New Zealand), Annika Hinze (University of Waikato, New Zealand), Nicholas Vanderschantz (University of Waikato, New Zealand), Claire Timpany (University of Waikato, New Zealand), and Sarah-Jane Saravani (University of Waikato, New Zealand)	266
Sharing is Caring! Joint Multitask Learning Helps Aspect-Category Extraction and Sentiment Detection in Scientific Peer Reviews Sandeep Kumar (Indian Institute of Technology Patna, India), Tirthankar Ghosal (Charles University, CZ), Prabhat Kumar Bharti (Indian Institute of Technology Patna, India), and Asif Ekbal (Indian Institute of Technology Patna, India)	. 270
Three Dimensions of Science: A Web Tool for 3D Visualization of Scientific Literature	274
JCDL 2021 Posters and Demos	
Academic Storage Cluster Alexander von Tottleben (University of Wuppertal, Germany), Cornelius Ihle (University of Wuppertal, Germany), Moritz Schubotz (FIZ – Karlsruhe, Germany), and Bela Gipp (University of Wuppertal, Germany)	278
ACM-CR: A Manually Annotated Test Collection for Citation Recommendation	280
Analyzing Unconstrained Reading Patterns of Digital Documents Using Eye Tracking	282
Are Altmetrics Proxies or Complements to Citations for Assessing Impact in Computer Science? Yusra Shakeel (Otto-von-Guericke University Magdeburg, Germany), Rand Alchokr (Otto-von-Guericke University Magdeburg, Germany), Jacob Krüger (Otto-von-Guericke University Magdeburg, Germany; Ruhr-University Bochum, Germany), Gunter Saake (Otto-von-Guericke University Magdeburg, Germany), and Thomas Leich (Harz University Wernigerode, Germany)	. 284
Assisted Text Annotation Using Active Learning to Achieve High Quality with Little Effort Franziska Weeber (University of Konstanz, Germany), Felix Hamborg (Heidelberg Academy of Sciences and Humanities, Germany), Karsten Donnay (Heidelberg Academy of Sciences and Humanities, Germany; University of Zurich, Switzerland), and Bela Gipp (University of Wuppertal, Germany)	287
Automatic Recognition of Learning Resource Category in a Digital Library	289

Building the COVID-19 Portal By Integrating Literature, Clinical Trials, and Knowledge Graphs	291
Aosen Wan (University of Texas at Austin), Yan Zhan (University of Texas at Austin), Sanjana Tripathi (University of Texas at Austin), Jiarong Yang (University of Texas at Austin), Mona Sachdev (University of Texas at Austin), Shreya Paithankar (Michigan State University), Joel Duerksen (Data2Discovery), Bin Chen (Michigan State University), and Ying Ding (University of Texas at Austin)	
Cell Block HTML: Towards Spreadsheet-Based Text-Mining for the Masses	293
ConSTR: A Contextual Search Term Recommender Thomas Krämer (GESIS – Leibniz Institute for the Social Sciences, Germany), Zeljko Carevic (GESIS – Leibniz Institute for the Social Sciences, Germany), Dwaipayan Roy (IISER, India), Claus-Peter Klas (GESIS – Leibniz Institute for the Social Sciences, Germany), and Philipp Mayr (GESIS – Leibniz Institute for the Social Sciences, Germany)	295
Crowdsourced Linked Data Question Answering with AQUACOLD	297
DSDB: An Open-Source System for Database Versioning & Curation	299
Evaluating BERT's Encoding of Intrinsic Semantic Features of OCR'd Digital Library Collections Ming Jiang (University of Illinois at Urbana-Champaign, USA), Yuerong Hu (University of Illinois at Urbana-Champaign, USA), Glen Worthey (University of Illinois at Urbana-Champaign, USA), Ryan C Dubnicek (University of Illinois at Urbana-Champaign, USA), Ted Underwood (University of Illinois at Urbana-Champaign, USA), and J Stephen Downie (University of Illinois at Urbana-Champaign, USA)	308
Extending Chromium: Memento-Aware Browser	310
Finding the Relevance Between Publication Venues Based on Research Trend Similarity and Citation Relationships	312
Girl with a Pearl Earring: Supporting 'Close Reading' of Art in a Digital Library	314
Hypercane: Intelligent Sampling for Web Archive Collections	316

Incorporating Fairness in Paper Recommendation Reem Alsaffar (University of Arkansas, USA), Susan Gauch (University of Arkansas, USA), Mohammed Alqahtani (University of Arkansas, USA), and Omar Salman (University of Arkansas, USA)	318
Sharing the Past: the Library as Digital Co-Design Space for Intergenerational Heritage Preservation	320
Peter J. Cobb (University of Hong Kong), Esther M. W. Woo (University of Hong Kong), Nicol F. C. Pan (University of Hong Kong), Vivian W. Q. Lou (University of Hong Kong), Xiao Hu (University of Hong Kong), Michael Cheng (University of Hong Kong), and Jesse Xiao (University of Hong Kong)	
TASSY—A Text Annotation Survey System	322
Towards A Reliable Ground-Truth For Biased Language Detection Timo Spinde (University of Konstanz, Germany), David Krieger (University of Konstanz, Germany), Manuel Plank (University of Konstanz, Germany), and Bela Gipp (University of Wuppertal, Germany)	324
Towards Transparent Data Cleaning: The Data Cleaning Model Explorer (DCM/X)	326
TweetPap: A Dataset to Study the Social Media Discourse of Scientific Papers	328
Understanding the Contributions of Junior Researchers at Software-Engineering Conferences Rand Alchokr (Otto von Guericke University Magdeburg), Jacob Krüger (Ruhr University Bochum), Yusra Shakeel (Otto von Guericke University Magdeburg), Gunter Saake (Otto von Guericke University Magdeburg), and Thomas Leich (Harz University Wernigerode)	330
User-Centred Application for Modeling Journeys in Digital Libraries Marwa Trabelsi (University of La Rochelle, France), Cyrille Suire (University of La Rochelle, France), Jacques Morcos (University of La Rochelle, France), and Ronan Champagnat (University of La Rochelle, France)	332
VeTo-web: A Recommendation Tool for the Expansion of Sets of Scholars	334
Visualizing the Evolution of Information Retrieval via the ACM Computer Classification Codes	336
Jan W. Buzydlowski (n/a) and Lillian N. Cassel (Villanova University)	550

Weak Supervision for Scientific Document Relevance Tagging Drahomira Herrmannova (Oak Ridge National Laboratory, USA), Chathika Gunaratne (Oak Ridge National Laboratory, USA), Vickie Walker (National Institute of Environmental Health Sciences, USA), Andrew Rooney (National Institute of Environmental Health Sciences, USA), Robert Patton (Oak Ridge National Laboratory, USA), Mary Wolfe (National Institute of Environmental Health Sciences, USA), and Charles Schmitt (National Institute of Environmental Health Sciences, USA)	. 338
What Did It Look Like: A Service for Creating Website Timelapses using the Memento Framework Dhruv Patel (Old Dominion University, USA), Michele C. Weigle (Old Dominion University, USA), Michael L. Nelson (Old Dominion University, USA), and Alexander C. Nwala (Indiana University, USA)	340
What Were People Searching For? A Query Log Analysis of An Academic Search Engine	. 342
JCDL 2021: Workshops	
1st International Workshop on Digital Language Archives	344
Digital Infrastructures for Scholarly Content Objects	. 346
HistoInformatics2021: The 6th International Workshop on Computational History Yasunobu Sumikawa (Takushoku University, Japan), Ryohei Ikejiri (The University of Tokyo, Japan), Antoine Doucet (University of La Rochelle, France), Eva Pfanzelter (University of Innsbruck, Austria), Mohammed Hasanuzzaman (Technological University, Ireland), Gaël Dias (University of Caen Normandie, France), Ian Milligan (University of Waterloo, Canada), and Adam Jatowt (University of Innsbruck, Austria)	348
Workshop on the Future of Digital Libraries	350
JCDL 2021: Tutorials	
Introduction to and Hands-On Use Cases with HathiTrust Research Center's Extracted Features 2.0 Dataset	352

Introduction to Digital Libraries	354
Edward A. Fox (Virginia Tech, USA) and Yinlin Chen (Virginia Tech, USA)	
JCDL 2021 Tutorial on Systemic Challenges and Computational Solutions on Bias and	
Unfairness in Peer Review	356
Nihar B. Shah (Carnegie Mellon University, USA)	
Author Index	359