

# **2025 ACM/IEEE Joint Conference on Digital Libraries (JCDL 2025)**

**Dekalb, Illinois, USA  
15-19 December 2025**



**IEEE Catalog Number: CFP25JCD-POD  
ISBN: 979-8-3315-6804-7**

**Copyright © 2025 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:	CFP25JCD-POD
ISBN (Print-On-Demand):	979-8-3315-6804-7
ISBN (Online):	979-8-3315-6803-0

**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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2025 ACM/IEEE Joint Conference on Digital Libraries (JCDL) JCDL 2025

## Table of Contents

Message from the General Chair .....	xii
Message from the Program Committee Chairs .....	xiv
Organizing Committee .....	xv
Program Committee .....	xvii
Steering Committee .....	xxii
Keynotes .....	xxiii
Sponsors .....	xxvi

## Long Papers

SciCom Wiki: A Digital Library to Support the Science Communication Knowledge Infrastructure for Videos and Podcasts .....	1
<i>Tim Wittenborg (Leibniz University Hannover, Germany), Niklas Stehr (Leibniz University Hannover, Germany), Oliver Karras (TIB - Leibniz Information Centre for Science and Technology, Germany), and Sören Auer (Leibniz University Hannover, Germany; TIB - Leibniz Information Centre for Science and Technology, Germany)</i>	
How a Tortured Conference Becomes a Series: An Analysis of Conference Manipulations .....	11
<i>Wendeline Swart (University of Toulouse, France), Guillaume Cabanac (University of Toulouse, France; Institut Universitaire de France (IUF), France), Ophélie Fraisier-Vannier (University of Toulouse, France), and Gilles Hubert (University of Toulouse, France)</i>	
CC30k: A Citation Contexts Dataset for Reproducibility-Oriented Sentiment Analysis .....	20
<i>Rochana R. Obadage (Old Dominion University, USA), Sarah Rajtmajer (The Pennsylvania State University, University Park, USA), and Jian Wu (Old Dominion University, USA)</i>	
Beyond CER and WER : How Does OCR Really Impact Information Retrieval? .....	30
<i>Alexandre Jaud (University of La Rochelle, France), Ahmed Hamdi (University of La Rochelle, France), Antoine Doucet (University of La Rochelle, France; University of Ljubljana, Slovenia), Adam Jatowt (University of Innsbruck, Austria), and Mickaël Coustaty (University of La Rochelle, France)</i>	
LLM-Generated Description and Reasoning: Use-Case for Library Recommendations .....	40
<i>Arash Sal Moslehian (University of Lausanne, Switzerland), Eya Briki (EPFL, Switzerland), and Michalis Vlachos (University of Lausanne, Switzerland)</i>	

Adaptive Progressive Fine-Tuning of VLMs for Long-Tailed Multimodal Retrieval .....	50
<i>Farid Alijani (Tampere University, Finland), Elina Late (Tampere University, Finland), and Sanna Kumpulainen (Tampere University, Finland)</i>	
Resilience, Volume, and Temporal Trends Across 25 Years of the Wayback Machine .....	59
<i>Kritika Garg (Old Dominion University, USA), Sawood Alam (Internet Archive, USA), Dietrich Ayala (Filecoin Foundation, Netherlands), Michele C. Weigle (Old Dominion University, USA), and Michael L. Nelson (Old Dominion University, USA)</i>	
The Indispensable Ren: A Multi-Method Study of Confucian Ethical Configurations in Trolling with Misinformation .....	69
<i>Xiao-Liang Shen (Wuhan University, China), Qianwen Qian (Wuhan University, China), and Jian Mou (Pusan National University, South Korea)</i>	
How to Evaluate the Carbon Footprint of Digital Libraries Services: The Case of Citation Verification .....	79
<i>Aysegul DEMIR (Univ. Grenoble Alpes, France), Qinyue LIU (Univ. Grenoble Alpes, France), and Cyril Labbé (Univ. Grenoble Alpes, France)</i>	
ARTO: An Artwork Object Ontology for Descriptive and Contextual Representation .....	89
<i>Can Yang (Australian National University, Australia), Bernardo Pereira Nunes (Australian National University, Australia), Sergio Rodríguez Méndez (Australian National University, Australia), Yige Chen (Australian National University, Australia), Rubén Manrique (Universidad de los Andes, Colombia), and Marco Antonio Casanova (Pontifical Catholic University of Rio de Janeiro, Brazil)</i>	
Scaffolding Inquiry-Oriented Web Search using LLM-Based Question Generation .....	99
<i>Yusuke Yamamoto (Nagoya City University, Japan)</i>	
Multi-Disciplinary Dataset Discovery from Citation-Verified Literature Contexts .....	109
<i>Zhiyin Tan (Leibniz University Hannover, Germany) and Changxu Duan (Technische Universität Darmstadt, Germany)</i>	
Screening Crossref for Integrity Issues: Massive and Longitudinal Mining of Publication Metadata .....	119
<i>Jules di Scala (Université de Toulouse, France), Guillaume Cabanac (Université de Toulouse, France; Institut Universitaire de France (IUF), France), Ophélie Fraisier-Vannier (Université de Toulouse, France), and Ronan Tournier (Université Toulouse 1 Capitole, France)</i>	
Learning from LLM Disagreement in Retrieval Evaluation .....	129
<i>William A. Ingram (University Libraries, USA), Bipasha Banerjee (University Libraries, USA), and Edward A. Fox (Virginia Tech, USA)</i>	
A Multi-Agent System for Complex Reasoning in Radiology Visual Question Answering .....	139
<i>Ziruo Yi (University of North Texas, USA), Jinyu Liu (University of North Texas, USA), Mark V. Albert (University of North Texas, USA), and Ting Xiao (University of North Texas, USA)</i>	

Generation, Evaluation, and Explanation of Novelists' Styles with Single-Token Prompts .....	148
<i>Mosab Rezaei (Northern Illinois University), Mina Rajaei Moghadam (Northern Illinois University), Abdul Rahman Shaikh (Northern Illinois University), Hamed Alhoori (Northern Illinois University), and Reva Freedman (Northern Illinois University)</i>	
CiteScreener: A Pipeline for Citation Verification in Digital Libraries with Datasets .....	158
<i>Qinyue Liu (Univ. Grenoble Alpes, France), Yağmur Öztürk (Univ. Grenoble Alpes, France), Tiziri Terkmani (Univ. Grenoble Alpes, France), François Portet (Univ. Grenoble Alpes, France), and Cyril Labbé (Univ. Grenoble Alpes, France)</i>	
A Self-Questioning Framework Towards Knowledge Self-Organization in Children's Readings via Prompt Learning and Fine-Tuning .....	167
<i>Jiacheng Yao (East China Normal University, China), Guoxiu He (East China Normal University, China), and Xin Xu (East China Normal University, China)</i>	
Identifying Future Work Chapters in Electronic Theses and Dissertations .....	177
<i>Amr Aboelnaga (Virginia Tech, USA), Hajra Klair (Virginia Tech, USA), Hoda Eldardiry (Virginia Tech, USA), and William A. Ingram (University Libraries, USA)</i>	
SciKGDash: The Scientific Knowledge Graph Dashboard for Supporting Knowledge Curation .....	187
<i>Lena John (TIB - Leibniz Information Centre for Science and Technology, Germany), Sören Auer (TIB - Leibniz Information Centre for Science and Technology, Germany), and Oliver Karras (TIB - Leibniz Information Centre for Science and Technology, Germany)</i>	
Context-Based URL Classification for Open Access Datasets and Software in Scholarly Documents .....	197
<i>Lamia Salsabil (Old Dominion University, USA), Rochana R. Obadage (Old Dominion University, USA), Bipasha Banerjee (Virginia Polytechnic Institute and State University, USA), Yasasi Abeysinghe (Old Dominion University, USA), Sawood Alam (Internet Archive, USA), Michael Färber (ScaDS.AI, Germany), William Ingram (Virginia Polytechnic Institute and State University, USA), Edward Fox (Virginia Polytechnic Institute and State University, USA), and Jian Wu (Old Dominion University, USA)</i>	
ReviewGuard: Enhancing Deficient Peer Review Detection via LLM-Driven Data Augmentation ...	207
<i>Haoxuan Zhang (University of North Texas, USA), Ruochi Li (North Carolina State University, USA), Sarthak Shrestha (University of North Texas, USA), Shree Harshini Mamidala (University of North Texas, USA), Revanth Putta (University of North Texas, USA), Arka Krishan Aggarwal (University of North Texas, USA), Ting Xiao (University of North Texas, USA), Junhua Ding (University of North Texas, USA), and Haihua Chen (University of North Texas, USA)</i>	
Linking Policies for the Permanent Web .....	217
<i>Patrick Le (Drexel University, USA), Quang Bui (Drexel University, USA), Alexander Grigorian (Drexel University, USA), Alexey Kuraev (Drexel University, USA), Thiyazan Salman (Drexel University, USA), Tu H. Nguyen (Drexel University, USA), Mat Kelly (Drexel University, USA), and Sawood Alam (Internet Archive, USA)</i>	

Age-Specific Fine-Tuning of Large Language Models with Sentiment-Guided Emotion Vectors for Teen Book Recommendations .....	227
<i>Yiu-Kai Ng (Brigham Young University, USA)</i>	

## Short Papers

How Devices Shape Mental Effort in Digital Document Reading: An Eye-Tracking Study .....	237
<i>Kumushini Thennakoon (Old Dominion University, USA), Yasasi Abeyasinghe (Old Dominion University, USA), Pasindu Thenahandi (Old Dominion University, USA), Lawrence Obiuvewwi (Old Dominion University, USA), Vikas Ashok (Old Dominion University, USA), and Sampath Jayarathna (Old Dominion University, USA)</i>	
NIED: A Corpus for Numeric Information Extraction from Dataset Descriptions .....	241
<i>Moriyuki Kamoto (Doshisha University, Japan), Akihiro Tamura (Doshisha University, Japan), and Marie Katsurai (Doshisha University, Japan)</i>	
Detecting Framing Bias in News via Probabilistic Graphical Modeling .....	245
<i>Shirin Shujaa (RMIT University, Vietnam) and Ginel Dorleon (RMIT University, Vietnam)</i>	
OntExtract: PROV-O Provenance Tracking for Document Analysis Workflows .....	249
<i>Christopher B. Rauch (Drexel University, USA), Hyung Wook Choi (Drexel University, USA), and Mat Kelly (Drexel University, USA)</i>	
K-Span Select and Multi-Dimensional Judging for Reliable Scholarly Question Answering .....	253
<i>Preetam Pati (Indian Institute of Information Technology Kalyani, India), Sayan De (Indian Institute of Information Technology Kalyani, India; Dr. B. C. Roy Engineering College, India), Saurabh Tiwari (Indian Institute of Information Technology Kalyani, India), Imon Mukherjee (Indian Institute of Information Technology Kalyani, India), and Debarshi Kumar Sanyal (Indian Association for the Cultivation of Science, India)</i>	
What Lies Beneath: A Call for Distribution-Based Visual Question & Answer Datasets .....	257
<i>Jill Naiman (University of Illinois Urbana-Champaign, USA), Daniel J. Evans (University of Illinois Urbana-Champaign, USA), and JooYoung Seo (University of Illinois Urbana-Champaign, USA)</i>	
Identifying and Classifying Software Mentions in full Text Scholarly Documents .....	261
<i>David Pride (The Open University, UK), Matteo Guenci (University of Bologna, Italy), Martin Docekal (Brno University of Technology, Czechia), Silvio Peroni (University of Bologna, Italy), and Petr Knoth (The Open University, UK)</i>	
Efficient Citation Screening by Weak Classifier Ensemble .....	265
<i>Xiaorui Jiang (University of Sheffield, Sheffield), Opeoluwa Akinseloyin (Coventry University, Coventry), and Vasile Palade (Coventry University, Coventry)</i>	
Better by Comparison: Retrieval-Augmented Contrastive Reasoning for Automatic Prompt Optimization .....	269
<i>Juhyeon Lee (Peking University, China), Wonduk Seo (Enhans, South Korea), Hyunjin An (Enhans, South Korea), Seunghyun Lee (Enhans, South Korea), and Yi Bu (Peking University, China)</i>	

Pre-Crawl Prioritization and Seed Classification for Large-Scale Web Archiving .....	273
<i>Akshith Garapati (Purdue University, USA) and Sawood Alam (Internet Archive, USA)</i>	

## Resources (Data/Software)

A Longitudinal Dataset of URLs Sampled From the Wayback Machine .....	277
<i>Kritika Garg (Old Dominion University, USA), Sawood Alam (Internet Archive, USA), Dietrich Ayala (Filecoin Foundation, Netherlands), Michele C. Weigle (Old Dominion University, USA), and Michael L. Nelson (Old Dominion University, USA)</i>	
Book Reconciler: An Open-Source Tool for Metadata Enrichment and Work-Level Clustering .....	281
<i>Matthew Miller (Library of Congress, USA), Dan Sinykin (Emory University, USA), and Melanie Walsh (University of Washington Information School, USA)</i>	
Fact-Stance: A Stance-Aware Dataset of Structured Scientific Claims with LLM Annotators .....	285
<i>Xin Lin (University of Chinese Academy of Sciences, China), Yang Zhao (University of Chinese Academy of Sciences, China), Zhixiong Zhang (University of Chinese Academy of Sciences, China), Yajiao Wang (University of Chinese Academy of Sciences, China), Yang Li (University of Chinese Academy of Sciences, China), and Mengting Zhang (University of Chinese Academy of Sciences, China)</i>	
PaperCross: A Cross-Document and Multi-Modal Question Answering Benchmark for Scientific Papers .....	289
<i>Guangyin Zhang (Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Liping Gu (Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Yang Li (Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Meng Wang (Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), and Mengting Zhang (Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)</i>	
End of Term Web Archive Dataset: Documenting the 2024 Presidential Transition .....	293
<i>Mark E. Phillips (University of North Texas, USA), Kristy K. Phillips (University of North Texas, USA), and Sawood Alam (Internet Archive, USA)</i>	
MAT-VB: MATHematical Text-Vision Benchmark .....	297
<i>Behrooz Mansouri (University of Southern Maine, USA), Aidan Bell (University of Southern Maine, USA), Nicholas Largey (University of Southern Maine, USA), and Abigail Pitcairn (University of Southern Maine, USA)</i>	

## Doctoral Consortium

The Ethical Duality of Human-AI Interaction:HowHuman and AI Ethics Shape CollaborativeRumorCombating Behavior .....	301
<i>Qianwen Qian (Wuhan University, China)</i>	

Fostering Knowledge Infrastructures in Science Communication and Aerospace Engineering .....	305
<i>Tim Wittenborg (Leibniz University Hannover, Germany)</i>	
AI Agents as Knowledge Explorers: A Case Study of 398 U.S. Case Law Quotations .....	309
<i>Jeremiah Milbauer (Carnegie Mellon University)</i>	
Equipping a Virtual Reading Promoter for Digital Libraries with Retrieval-Augmented Generation .....	313
<i>Zhenyu Li (Shanghai University, China)</i>	

## Posters / Demos

Multimodal Emotion Classification in Artwork: A Comparative Study Across Modalities .....	316
<i>Clayton Durepos (University of Southern Maine), Abigail Pitcairn (University of Southern Maine), and Behrooz Mansouri (University of Southern Maine)</i>	
A Lexically-Driven Adaptive Validation Framework for Search-Based Reference Matching .....	320
<i>Katsuyuki Hirai (Niigata University of Health and Welfare, Japan), Teruhito Kanazawa (National Institute of Informatics, Japan), and Takahiro Hayashi (Kansai University, Japan)</i>	
LLM-Based Active Learning for Identifying References to Archival Repositories .....	324
<i>Tokinori Suzuki (University of Tsukuba)</i>	
From Speech to LaTeX: Large Language Models for Mathematical Accessibility in Digital Libraries .....	328
<i>Abigail Pitcairn (University of Southern Maine, USA), Clayton Durepos (University of Southern Maine, USA), Nicholas Largey (University of Southern Maine, USA), and Behrooz Mansouri (University of Southern Maine, USA)</i>	
Revisiting the Link Accessibility Problem in Scholarly Papers with PLoS ONE Papers .....	332
<i>Tyler Chen (Princess Anne High School, USA) and Jian Wu (Old Dominion University, USA)</i>	
Extending the Usability of Digital Music Libraries with Analytical Interfaces Case Study – Traditional Music Collections .....	336
<i>Anna Maria Matuszewska (Polish Academy of Sciences, Poland)</i>	
ReAct-ExtrAct: A Tool for Source-Grounded Automated Data Extraction in Systematic Reviews ...	340
<i>Sebastian Krawczyk (Georgia Institute of Technology, USA), Paweł Jemioło (AGH University of Krakow, Poland), Jan Karkowski (Cracow University of Technology, Poland), Ilinka Ivanoska (Ss. Cyril and Methodius University in Skopje, North Macedonia), Miroslav Mirchev (Complexity Science Hub, Austria), and Wojciech Kusa (NASK National Research Institute, Poland)</i>	
Preserving and Retrieving Hindustani Music Notation: A Symbolic Music Processing Architecture for Digital Libraries .....	344
<i>Chandan Misra (XIM University, India), Attreye Chakraborty (XIM University, India), and Shreya Dutta (XIM University, India)</i>	

MathMex-V2: A Large Language Model Enabled Math Search Engine .....	348
<i>Clayton Durepos (University of Southern Maine), Ian McLaughlin (University of Southern Maine), Connor Lund (Falmouth High School), Anthony Sienbenmorgen (University of Southern Maine), Nicholas Largey (University of Southern Maine), Abigail Pitcairn (University of Southern Maine), and Behrooz Mansouri (University of Southern Maine)</i>	

## Workshops & Tutorials

Smarter Extraction of ScholArly METadata using Knowledge Graphs and Language Models .....	352
<i>Muhammad Asif Suryani (GESIS – Leibniz-Institut für Sozialwissenschaften), Brigitte Mathiak (GESIS – Leibniz-Institut für Sozialwissenschaften), Florian Reitz (Schloss Dagstuhl Leibniz-Zentrum für Informatik), Florian Jäckel (Schloss Dagstuhl Leibniz-Zentrum für Informatik), and Ansgar Scherp (Data Science and Big Data Analytics, Ulm University)</i>	
The 2nd International Workshop on Artificial Intelligence for the Science of Science .....	355
<i>Jian Wu (Old Dominion University), Sarah Rajtmajer (Pennsylvania State University), Yian Yin (Cornell University), Yi He (College of William &amp; Mary), and Staša Milojević (Indiana University, Bloomington)</i>	
3rd International Workshop on Digital Language Archives .....	358
<i>Oksana L. Zavalina (University of North Texas, USA), Shobhana L. Chelliah (University of Indiana, USA), and Mary Burke (University of North Texas, USA)</i>	
BESIDES: Beyond Search Engines-Intelligent Knowledge Discovery From Scholarly Publications...	361
<i>Gautam Shahi (University of Duisburg-Essen, Germany ) and Oliver Hummel (Technical University of Mannheim)</i>	

## Panels

Application of GAI to Empirical Studies: It's High Time to Discuss Challenges and Pitfalls.....	364
<i>Miyoung Chong (University of South Florida, USA), Junhua Ding (University of North Texas, USA), Chirag Shah (University of Washington, USA), and Haihua Chen (University of North Texas, USA)</i>	
<b>Author Index</b> .....	<b>367</b>