

*Commenced Publication in 1973*

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

## Editorial Board

David Hutchison

*Lancaster University, Lancaster, UK*

Takeo Kanade

*Carnegie Mellon University, Pittsburgh, PA, USA*

Josef Kittler

*University of Surrey, Guildford, UK*

Jon M. Kleinberg

*Cornell University, Ithaca, NY, USA*

Friedemann Mattern

*ETH Zurich, Zurich, Switzerland*

John C. Mitchell

*Stanford University, Stanford, CA, USA*

Moni Naor

*Weizmann Institute of Science, Rehovot, Israel*

C. Pandu Rangan

*Indian Institute of Technology Madras, Chennai, India*

Bernhard Steffen

*TU Dortmund University, Dortmund, Germany*

Demetri Terzopoulos

*University of California, Los Angeles, CA, USA*

Doug Tygar

*University of California, Berkeley, CA, USA*

Gerhard Weikum

*Max Planck Institute for Informatics, Saarbrücken, Germany*

More information about this series at <http://www.springer.com/series/7409>


Milena Dobreva · Annika Hinze  
Maja Žumer (Eds.)


# Maturity and Innovation in Digital Libraries

20th International Conference  
on Asia-Pacific Digital Libraries, ICADL 2018  
Hamilton, New Zealand, November 19–22, 2018  
Proceedings

*Editors*

Milena Dobrevá   
University College London Qatar  
Doha, Qatar

Maja Žumer   
University of Ljubljana  
Ljubljana, Slovenia

Annika Hinze   
University of Waikato  
Hamilton, New Zealand

ISSN 0302-9743                      ISSN 1611-3349 (electronic)  
Lecture Notes in Computer Science  
ISBN 978-3-030-04256-1              ISBN 978-3-030-04257-8 (eBook)  
<https://doi.org/10.1007/978-3-030-04257-8>

Library of Congress Control Number: 2018960913

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

© Springer Nature Switzerland AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

## Preface by Program Chairs

This volume contains the papers presented at the 20th International Conference on Asia-Pacific Digital Libraries (ICADL 2018), held during November 19–22, 2018, in Hamilton New Zealand. Since its beginnings in Hong Kong in 1998, ICADL has become one of the premier international conferences for digital library research. The conference series explores digital libraries as a broad foundation for interaction with information and information management in the networked information society. Bringing together the achievements in digital libraries research and development from Asia and Oceania with work from around the globe, ICADL serves as a unique forum where the regional drive for exposure of the rich digital collections often requiring out-of-the box solutions meets global excellence.

ICADL 2018 at the University of Waikato in New Zealand offered a further valuable opportunity for researchers, educators, and practitioners to share their experiences and innovative developments. The conference was held at Waikato University in Hamilton, New Zealand, a city of 140,000 people centered on the Waikato River in the heart of New Zealand’s rolling pastures. The main theme of ICADL 2018 was “Maturity and Innovation in Digital Libraries.” We invited high-quality, original research papers as well as practitioner papers identifying research problems and future directions. Submissions that resonate with the conference’s theme were especially welcome, but all topics in digital libraries were given equal consideration.

After 20 years, digital libraries have certainly reached maturity. At first glance, research of digital libraries on the general level is no longer needed and the focus has moved to more specialized areas – but in fact many new challenges have arisen. The proliferation of social media, a renewed interest in cultural heritage, digital humanities, research data management, new privacy legislation, changing user information behavior are only some of the topics that require a thorough discussion and an interdisciplinary approach. ICADL remains an excellent opportunity to bring together expertise from the region and beyond to explore the challenges and foster innovation.

The 2018 ICADL conference was co-located with the annual meeting of the Asia-Pacific iSchools Consortium and brought together a diverse group of academic and professional community members from all parts of the world to exchange their knowledge, experience, and practices in digital libraries and other related fields.

ICADL 2018 received 77 submissions from over 20 countries, attracting also this year some submissions from the Middle East, a region that does not have its own designated digital library forum. Each paper was carefully reviewed by the Program Committee members. Finally, 20 full papers and six short papers were selected to be presented, and 11 submissions were invited as work-in-progress papers. The submissions to ICADL 2018 covered a wide spectrum of topics from various areas, including semantic modeling, social media, Web and news archiving, heritage and localization, user experience, digital library technology, digital library use cases, research data management, librarianship, and education.

On behalf of the Organizing and Program Committees of ICADL 2018, we would like to express our appreciation to all the authors and attendees for participating in the conference. We also thank the sponsors, Program Committee members, external reviewers, supporting organizations, and volunteers for making the conference a success. Without their efforts, the conference would not have been possible.

We also acknowledge the University of Waikato for hosting the conference and hope the following years of this forum bring further maturity and innovation drive to the global digital library community.

November 2018

Milena Dobрева  
Annika Hinze  
Maja Žumer

## **Preface by Steering Committee Chair**

The International Conference on Asia-Pacific Digital Libraries (ICADL) started in 1998 in Hong Kong. ICADL is widely recognized as a top-level international conference similar to conferences such as JCDL and TPDL (formerly ECDL), although geographically based in the Asia-Pacific region. The conference has been successful in not only collecting novel research results, but also connecting people across global and regional communities. The Asia-Pacific region has significant diversity in many aspects such as culture, language, and development levels, and ICADL has been successful in bringing novel technologies and ideas to the regional communities and connecting their members.

In its 20-year history, ICADL has been held in 12 countries across the Asia-Pacific region. The 20th conference was hosted by the University of Waikato in New Zealand – a country that ICADL has never visited. The University of Waikato is well known in the digital library community as a leading institution from the early days of digital libraries research, e.g., Greenstone Digital Library.

As the chair of the Steering Committee of ICADL, I would like to thank the organizers of ICADL 2018 for their tremendous efforts. I would also like to thank all the organizers of the past conferences who have contributed not only in keeping the quality of ICADL high but also in helping establish and expand the Asia-Pacific digital library research community. Last but not least, I would like to thank all the people who have contributed to ICADL as an author, a reviewer, a tutor, a workshop organizer, or audience member.

Shigeo Sugimoto

# Organization

## Program Committee Co-chairs

Annika Hinze	University of Waikato, New Zealand
Maja Žumer	University of Ljubljana, Slovenia
Milena Dobrova	UCL Qatar, Qatar

## General Conference Chairs

Sally Jo Cunningham	University of Waikato, New Zealand
David Bainbridge	University of Waikato, New Zealand

## Workshop and Tutorial Chair

Trond Aalberg	Norwegian University of Science and Technology, Norway
---------------	---

## Poster and Work-in-Progress Chair

Fernando Loizides	Cardiff University, UK
-------------------	------------------------

## Doctoral Consortium Chairs

Sally Jo Cunningham	University of Waikato, New Zealand
Nicholas Vanderschantz	University of Waikato, New Zealand

## Treasurer

David Nichols	University of Waikato, New Zealand
---------------	------------------------------------

## Local Organization

Renee Hoad  
Tania Robinson  
Bronwyn Webster

## Program Committee and Reviewers

Trond Aalberg	Norwegian University of Science and Technology, Norway
Sultan Al-Daihani	University of Kuwait, Kuwait
Mohammad Aliannejadi	University of Lugano (USI), Switzerland



Seyed Ali Bahrainian	University of Lugano (USI), Switzerland
Maram Barifah	University of Lugano (USI), Switzerland
Erik Champion	Curtin University, Australia
Manajit Chakraborty	University of Lugano (USI), Switzerland
Songphan Choemprayong	Chulalongkorn University, Thailand
Gobinda Chowdhury	Northumbria University, UK
Fabio Crestani	University of Lugano (USI), Switzerland
Schubert Foo	Nanyang Technological University, Singapore
Edward Fox	Virginia Tech, USA
Dion Goh	Nanyang Technological University, Singapore
Yunhyong Kim	University of Glasgow, UK
Adam Jatowt	Kyoto University, Japan
Wachiraporn Klungthanaboon	Chulalongkorn University, Thailand
Chern Li Liew	Victoria University of Wellington, New Zealand
Fernando Loizides	Cardiff University, UK
Naya Sucha-Xaya	Chulalongkorn University, Thailand
Esteban Andrés Ríssola	University of Lugano (USI), Switzerland
Akira Maeda	Ritsumeikan University, Japan
Tanja Mercun	University of Ljubljana, Slovenia
Francesca Morselli	DANS-KNAW
David Nichols	University of Waikato, New Zealand
Gillian Oliver	Monash University, Australia
Georgios Papaioannou	UCL Qatar, Qatar
Christos Papatheodorou	Ionian University, Greece
Jan Pisanski	University of Ljubljana, Slovenia
Edie Rasmussen	University of British Columbia, Canada
Andreas Rauber	Vienna University of Technology, Austria
Alicia Salaz	Carnegie Mellon University, Qatar
Ali Shiri	University of Alberta, Canada
Armin Straube	UCL Qatar, Qatar
Shigeo Sugimoto	University of Tsukuba, Japan
Giannis Tsakonas	University of Patras, Greece
Nicholas Vanderschantz	University of Waikato, New Zealand
Diane Velasquez	University of South Australia, Australia
Dan Wu	Wuhan University, China

# Contents

## Topic Modelling and Semantic Analysis

Evaluating the Impact of OCR Errors on Topic Modeling . . . . .	3
<i>Stephen Mutuvi, Antoine Doucet, Moses Odeo, and Adam Jatowt</i>	
Measuring the Semantic World – How to Map Meaning to High-Dimensional Entity Clusters in PubMed? . . . . .	15
<i>Janus Wawrzinek and Wolf-Tilo Balke</i>	
Towards Semantic Quality Enhancement of User Generated Content . . . . .	28
<i>José María González Pinto, Niklas Kiehne, and Wolf-Tilo Balke</i>	
Query-Based Versus Resource-Based Cache Strategies in Tag-Based Browsing Systems. . . . .	41
<i>Joaquín Gayoso-Cabada, Mercedes Gómez-Albarrán, and José-Luis Sierra</i>	
Automatic Labanotation Generation, Semi-automatic Semantic Annotation and Retrieval of Recorded Videos. . . . .	55
<i>Swati Dewan, Shubham Agarwal, and Navjyoti Singh</i>	
Quality Classification of Scientific Publications Using Hybrid Summarization Model . . . . .	61
<i>Hafiz Ahmad Awais Chaudhary, Saeed-Ul Hassan, Naif Radi Aljohani, and Ali Daud</i>	

## Social Media, Web, and News

Investigating the Characteristics and Research Impact of Sentiments in Tweets with Links to Computer Science Research Papers. . . . .	71
<i>Aravind Sesagiri Raamkumar, Savitha Ganesan, Keerthana Jothiramalingam, Muthu Kumaran Selva, Mojisola Erdt, and Yin-Leng Theng</i>	
Predicting Social News Use: The Effect of Gratifications, Social Presence, and Information Control Affordances. . . . .	83
<i>Winston Jin Song Teo</i>	
Aspect-Based Sentiment Analysis of Nuclear Energy Tweets with Attentive Deep Neural Network . . . . .	99
<i>Zhengyuan Liu and Jin-Cheon Na</i>	

Where the Dead Blogs Are: A Disaggregated Exploration of Web Archives to Reveal Extinct Online Collectives . . . . .	112
<i>Quentin Lobbé</i>	
A Method for Retrieval of Tweets About Hospital Patient Experience . . . . .	124
<i>Julie Walters and Gobinda Chowdhury</i>	
Rewarding, But Not for Everyone: Interaction Acts and Perceived Post Quality on Social Q&A Sites . . . . .	136
<i>Sei-Ching Joanna Sin, Chei Sian Lee, and Xinran Chen</i>	
Towards Recommending Interesting Content in News Archives . . . . .	142
<i>I-Chen Hung, Michael Färber, and Adam Jatowt</i>	
A Twitter-Based Culture Visualization System by Analyzing Multilingual Geo-Tagged Tweets. . . . .	147
<i>Yuanyuan Wang, Panote Siriaraya, Yusuke Nakaoka, Haruka Sakata, Yukiko Kawai, and Toyokazu Akiyama</i>	
<b>Heritage and Localization</b>	
Development of Content-Based Metadata Scheme of Classical Poetry in Thai National Historical Corpus . . . . .	153
<i>Songphan Choemprayong, Pittayawat Pittayaporn, Vipas Pothipath, Thaneerat Jatuthasri, and Jinawat Kaenmuang</i>	
Research Data Management by Academics and Researchers: Perceptions, Knowledge and Practices . . . . .	166
<i>Shaheen Majid, Schubert Foo, and Xue Zhang</i>	
Examining Japanese American Digital Library Collections with an Ethnographic Lens. . . . .	179
<i>Andrew B. Wertheimer and Noriko Asato</i>	
Exploring Information Needs and Search Behaviour of Swahili Speakers in Tanzania . . . . .	185
<i>Joseph P. Telemala and Hussein Suleman</i>	
Bilingual Qatar Digital Library: Benefits and Challenges . . . . .	191
<i>Mahmoud Sayed A. Mahmoud and Maha M. Al-Sarraj</i>	
Digital Preservation Effort of Manuscripts Collection: Case Studies of <i>pustakabudaya.id</i> as Indonesia Heritage Digital Library . . . . .	195
<i>Revi Kuswara</i>	

The General Data Protection Regulation (GDPR, 2016/679/EE) and the (Big) Personal Data in Cultural Institutions: Thoughts on the GDPR Compliance Process. . . . . 201

*Georgios Papaioannou and Ioannis Sarakinos*

A Recommender System in Ukiyo-e Digital Archives for Japanese Art Novices . . . . . 205

*Jiayun Wang, Biligsaikhan Batjargal, Akira Maeda, and Kyoji Kawagoe*

**User Experience**

Tertiary Students’ Preferences for Library Search Results Pages on a Mobile Device. . . . . 213

*Nicholas Vanderschantz, Claire Timpany, and Chun Feng*

Book Recommendation Beyond the Usual Suspects: Embedding Book Plots Together with Place and Time Information . . . . . 227

*Julian Risch, Samuele Garda, and Ralf Krestel*

Users’ Responses to Privacy Issues with the Connected Information Ecologies Created by Fitness Trackers . . . . . 240

*Zablon Pingo and Bhuva Narayan*

Multiple Level Enhancement of Children’s Picture Books with Augmented Reality . . . . . 256

*Nicholas Vanderschantz, Annika Hinze, and Aysha AL-Hashami*

A Visual Content Analysis of Thai Government’s Census Infographics . . . . . 261

*Somsak Sriborisutsakul, Sorakom Dissamana, and Saowapha Limwichitr*

**Digital Library Technology**

BitView: Using Blockchain Technology to Validate and Diffuse Global Usage Data for Academic Publications. . . . . 267

*Camillo Lamanna and Manfredi La Manna*

Adaptive Edit-Distance and Regression Approach for Post-OCR Text Correction. . . . . 278

*Thi-Tuyet-Hai Nguyen, Mickael Coustaty, Antoine Doucet, Adam Jatowt, and Nhu-Van Nguyen*

Performance Comparison of Ad-Hoc Retrieval Models over Full-Text vs. Titles of Documents . . . . . 290

*Ahmed Saleh, Tilman Beck, Lukas Galke, and Ansgar Scherp*

Acquiring Metadata to Support Biographies of Museum Artefacts . . . . .	304
<i>Can Zhao, Michael B. Twidale, and David M. Nichols</i>	
Mining the Context of Citations in Scientific Publications . . . . .	316
<i>Saeed-Ul Hassan, Sehrish Iqbal, Mubashir Imran, Naif Radi Aljohani, and Raheel Nawaz</i>	
A Metadata Extractor for Books in a Digital Library . . . . .	323
<i>Sk. Simran Akhtar, Debarshi Kumar Sanyal, Samiran Chattopadhyay, Plaban Kumar Bhowmick, and Partha Pratim Das</i>	
Ownership Stamp Character Recognition System Based on Ancient Character Typeface . . . . .	328
<i>Kangying Li, Biligsaikhan Batjargal, and Akira Maeda</i>	
<b>Use Cases and Digital Librarianship</b>	
Identifying Design Requirements of a User-Centered Research Data Management System . . . . .	335
<i>Maryam Bugaje and Gobinda Chowdhury</i>	
Going Beyond Technical Requirements: The Call for a More Interdisciplinary Curriculum for Educating Digital Librarians . . . . .	348
<i>Andrew Wertheimer and Noriko Asato</i>	
ETDs, Research Data and More: The Current Status at Nanyang Technological University Singapore. . . . .	356
<i>Schubert Foo, Xue Zhang, Yang Ruan, and Su Nee Goh</i>	
<b>Author Index</b> . . . . .	363