

2018 IEEE 34th International Conference on Data Engineering Workshops (ICDEW 2018)

**Paris, France
16-20 April 2018**



**IEEE Catalog Number: CFP1845A-POD
ISBN: 978-1-5386-6307-3**

**Copyright © 2018 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:	CFP1845A-POD
ISBN (Print-On-Demand):	978-1-5386-6307-3
ISBN (Online):	978-1-5386-6306-6
ISSN:	1943-2895

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

CURRAN ASSOCIATES INC.
proceedings
.com

2018 IEEE 34th International Conference on Data Engineering Workshops **ICDEW 2018**

Table of Contents

Message from the ICDE 2018 Chairs .ix.....	
ICDE 2018 Workshops Organizing and Program Committees .xii.....	

Context in Analytics Workshop (CiA 2018)

Research Directions in Context Analytics (Research - Full Papers)

Context Analytics: Vision, Architecture, Opportunity .1.....	
<i>Eser Kandogan (IBM Almaden Reserach Center), Mary Roth (IBM Almaden Reserach Center), Ignacio Terrizzano (IBM Almaden Reserach Center), Ed Seabolt (IBM Almaden Reserach Center), Peter Schwarz (IBM Almaden Reserach Center), Harsha Krishnareddy (IBM Almaden Reserach Center), and Akshay Agarwal (IBM Almaden Reserach Center)</i>	
Contextually-Enriched Querying of Integrated Data Sources .9.....	
<i>Giacomo Cavallo (University of Torino), Francesco Di Mauro (University of Torino), Paolo Pasteris (University of Torino), Maria Luisa Sapino (University of Torino), and Kasim Selcuk Candan (Arizona State University)</i>	

Exploiting Context in Analytics (Vision - Short Papers)

Exploiting Functional Context in Biology: Reconsidering Classification of Bacterial Life .17.....	
<i>James H. Kaufman (IBM Almaden Research Center), Ed Seabolt (IBM Almaden Research Center), Mark Kunitomi (IBM Almaden Research Center), Akshay Agarwal (IBM Almaden Research Center), Kristen Beck (IBM Almaden Research Center), Harsha Krishnareddy (IBM Almaden Research Center), and Bart C. Weimer (University of California)</i>	
Contextual Data Cleaning .21.....	
<i>Morteza Alipour-Langouri (McMaster University), Zheng Zheng (McMaster University), Fei Chiang (McMaster University), Lukasz Golab (University of Waterloo), and Jaroslaw Szlichta (University of Ontario Institute of Technology)</i>	
Datasets for Context-Aware Recommender Systems: Current Context and Possible Directions .25.....	
<i>Sergio Ilarri (University of Zaragoza), Raquel Trillo-Lado (University of Zaragoza), and Ramon Hermoso (University of Zaragoza)</i>	

Semantics of Context (Invited Paper)

What Do We Mean When We Talk about Context? .29.....	
<i>Naresh Kumar Agarwal (Simmons School of Library and Information Science)</i>	

Data Engineering Methods in Precision Medicine (DEPM 2018)

Distributed Privacy-Preserving Record Linkage Using Pivot-Based Filter Techniques .33.....	
<i>Marcel Gladbach (University of Leipzig), Ziad Sehili (University of Leipzig), Thomas Kudrass (Leipzig University of Applied Science), Peter Christen (Australian National University), and Erhard Rahm (University of Leipzig)</i>	
Evaluating Text Analytic Frameworks for Mental Health Surveillance .39.....	
<i>Benjamin Mayer (Oak Ridge National Laboratory), Joshua Arnold (Oak Ridge National Laboratory), Edmon Begoli (Oak Ridge National Laboratory), Everett Rush (Oak Ridge National Laboratory), Michael Drewry (Oak Ridge National Laboratory), Kris Brown (University of Tennessee), Eduardo Ponce (University of Tennessee), and Sudarshan Srinivas (University of Tennessee)</i>	
Improving Quality of Observational Streaming Medical Data by Using Long Short-Term Memory Networks (LSTMs) .48.....	
<i>Michael Bowie (Oak Ridge National Laboratory), Edmon Begoli (Oak Ridge National Laboratory), and Byung Park (Oak Ridge National Laboratory)</i>	

Data Engineering Meets the Semantic Web Workshop (DESWeb 2018)

RDF Query Answering Using Apache Spark: Review and Assessment .54.....	
<i>Giannis Agathangelos (ICS-FORTH), Georgia Troullinou (ICS-FORTH), Haridimos Kondylakis (ICS-FORTH), Kostas Stefanidis (University of Tampere), and Dimitris Plexousakis (ICS-FORTH)</i>	
Detecting Search and Rescue Missions from AIS Data .60.....	
<i>Iraklis Varlamis (Harokopio University of Athens), Konstantinos Tserpes (Harokopio University of Athens), and Christos Sardianos (Harokopio University of Athens)</i>	
Quotient RDF Summaries Based on Type Hierarchies .66.....	
<i>Pawel Guzewicz (Universite Paris-Saclay) and Ioana Manolescu (Inria and Ecole Polytechnique)</i>	
Dynamic Service Detection for Automated Patient Selection for Study Recruitment Purposes .72.....	
<i>Efthymios Chondrogiannis (National Technical University of Athens), Vassiliki Andronikou (National Technical University of Athens), Efsthios Karanastasis (National Technical University of Athens), and Theodora Varvarigou (National Technical University of Athens)</i>	
Towards Artificial Priority Queues for Similarity Query Execution .78.....	
<i>Matej Antol (Masaryk University) and Vlastislav Dohnal (Masaryk University)</i>	

Scaling Up Schema Discovery for RDF Datasets .84.....	
<i>Redouane Bouhamoum (University of Versailles Saint-Quentin-en-Yvelines), Kenza Kellou-Menouer (University of Versailles Saint-Quentin-en-Yvelines), Stephane Lopes (University of Versailles Saint-Quentin-en-Yvelines), and Zoubida Kedad (University of Versailles Saint-Quentin-en-Yvelines)</i>	

Joint Workshop of HardBD (International Workshop on Big Data Management on Emerging Hardware) and Active (Workshop on Data Management on Virtualized Active Systems)

Session 1

Exploiting Automatic Vectorization to Employ SPMD on SIMD Registers .90.....	
<i>Stefan Sprenger (Humboldt-Universität zu Berlin), Steffen Zeuch (German Research Center for Artificial Intelligence), and Ulf Leser (Humboldt-Universität zu Berlin)</i>	
Conflict Detection-Based Run-Length Encoding - AVX-512 CD Instruction Set in Action .96.....	
<i>Annett Ungethum (TU Dresden), Johannes Pietrzyk (TU Dresden), Patrick Damme (TU Dresden), Dirk Habich (TU Dresden), and Wolfgang Lehner (TU Dresden)</i>	
Fused Table Scans: Combining AVX-512 and JIT to Double the Performance of Multi-Predicate Scans .102.....	
<i>Markus Dreseler (Hasso Plattner Institute), Jan Kossmann (Hasso Plattner Institute), Johannes Frohnhofer (Hasso Plattner Institute), Matthias Uflacker (Hasso Plattner Institute), and Hasso Plattner (Hasso Plattner Institute)</i>	
An NVM-Aware Storage Layout for Analytical Workloads .110.....	
<i>Gotze Philipp (TU Ilmenau), Baumann Stephan (Actian Germany GmbH), and Sattler Kai-Uwe (TU Ilmenau)</i>	

Session 2

Workload-Driven Horizontal Partitioning and Pruning for Large HTAP Systems .116.....	
<i>Martin Boissier (Hasso Plattner Institute) and Kurzynski Daniel (SAP SE)</i>	
Life Cycle of Transactional Data in In-memory Databases .122.....	
<i>Amit Pathak (SAP Labs), Aditya Gurajada (SAP Labs), and Pushkar Khadilkar (SAP Labs)</i>	
Towards Batch-Processing on Cold Storage Devices .134.....	
<i>Ali Hadian (Imperial College London) and Thomas Heinis (Imperial College London)</i>	
Efficient Stream Processing of Scientific Data .140.....	
<i>Thomas Lindemann (TU Dortmund University), Jonas Kauke (TU Dortmund University), and Jens Teubner (TU Dortmund University)</i>	

Data Engineering Meets Intelligent Food and Cooking Recipe Workshop 2018 (DECOR 2018)

Challenges and BlockChain-based cOllaborative iNtelligence on Food, blg Data: cooking rEcipe, multimedia Nutrition, and healthCare initiatiVE

Data on Composition of Slovenian Food .146.....	
<i>Mojca Korosec (University of Ljubljana) and Jasna Bertonec (University of Ljubljana)</i>	
Collaborative Intelligence and Decentralized Business Community Building–Potentials in Food/Nutrition Sector .150.....	
<i>Iztok Petek (Luxembourg Slovenian Business Club) and Natasa Zajec (Luxembourg Slovenian Business Club)</i>	

Computational Cooking

The CRWB RSbench: Towards a Cooking Recipe Benchmark Initiative .154.....	
<i>Andres Frederic (National Institute of Informatics)</i>	
Data-Driven Investigations of Culinary Patterns in Traditional Recipes Across the World .157.....	
<i>Ganesh Bagler (Indraprastha Institute of Information Technology) and Navjot Singh (Indraprastha Institute of Information Technology)</i>	
DiEvent: Towards an Automated Framework for Analyzing Dining Events .163.....	
<i>Mahmoud Qodseya (IRIT-Paul Sabatier University), Mahdi Washha (IRIT-Paul Sabatier University), and Florence Sedes (IRIT-Paul Sabatier University)</i>	
Images and Recipes: Retrieval in the Cooking Context .169.....	
<i>Micael Carvalho (Sorbonne Universite), Remi Cadene (Sorbonne Universite), David Picard (Sorbonne Universite), Laure Soulier (Sorbonne Universite), and Matthieu Cord (Sorbonne Universite)</i>	

Author Index .175	
-------------------------	--