

## Founding Editors

Gerhard Goos

*Karlsruhe Institute of Technology, Karlsruhe, Germany*

Juris Hartmanis

*Cornell University, Ithaca, NY, USA*


## Editorial Board Members

Elisa Bertino

*Purdue University, West Lafayette, IN, USA*

Wen Gao

*Peking University, Beijing, China*

Bernhard Steffen 

*TU Dortmund University, Dortmund, Germany*

Gerhard Woeginger 

*RWTH Aachen, Aachen, Germany*

Moti Yung

*Columbia University, New York, NY, USA*

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

Christine Strauss · Gabriele Kotsis ·  
A Min Tjoa · Ismail Khalil (Eds.)

# Database and Expert Systems Applications

32nd International Conference, DEXA 2021  
Virtual Event, September 27–30, 2021  
Proceedings, Part I

*Editors*

Christine Strauss  
University of Vienna  
Vienna, Austria

A Min Tjoa  
Vienna University of Technology  
Vienna, Austria

Gabriele Kotsis  
Johannes Kepler University of Linz  
Linz, Oberösterreich, Austria

Ismail Khalil  
Johannes Kepler University of Linz  
Linz, Austria

ISSN 0302-9743

ISSN 1611-3349 (electronic)

Lecture Notes in Computer Science

ISBN 978-3-030-86471-2

ISBN 978-3-030-86472-9 (eBook)

<https://doi.org/10.1007/978-3-030-86472-9>

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

© Springer Nature Switzerland AG 2021

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, expressed 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

The volume at hand represents the result of joint efforts of contributing researchers, reviewers, and organizers, and contains the papers presented at the 32nd International Conference on Database and Expert Systems Applications (DEXA 2021). This year, DEXA was held for the second time as a virtual conference during September 27–30, 2021, instead of in Linz, Austria, as originally planned. The decision to organize another virtual version of DEXA was driven by the intention to provide stable conditions for all DEXA participants and set a good example in temporarily suspending on-site meetings. We put our trust in the loyalty of DEXA community and look forward to personal DEXA meetings in 2022.

We are proud to report that authors from 43 different countries submitted papers to DEXA this year. The number of submissions was similar to those of the past few years. Our Program Committee conducted more than 500 reviews. We would like to sincerely thank our Program Committee members for their rigorous and critical, and at the same time motivating, reviews of these submissions. Based on the total number of accepted papers, we can report that the acceptance rate this year was 27%, a rate comparable to DEXA conferences of the last few years.

The conference program this year covered a wide range of important topics such as data management and analytics; consistency; integrity; quality of data; data analysis and data modeling; data mining; databases and data management; information retrieval; prediction and decision support; authenticity, privacy, security, and trust; cloud databases and workflows; data and information processing; knowledge discovery; machine learning; semantic web and ontologies; stream data processing; and temporal, spatial, and high dimensional databases.

We tried to follow our on-site face-to-face format. Thus, the authors of the accepted papers presented their research online using video conference software over four days. Presentations were performed live in 12 different thematic clusters structured as 15 sessions, each one with an assigned session chair. The scientific presentations, discussions, and question-and-answer time were all live and part of each session. As we were aware of time difference issues, for example, for participants from Australia or South American countries having to present or participate during unusual times of the day, we tried to minimize this inconvenience.

We would like to express our gratitude to the distinguished keynote speakers for illuminating us on their leading-edge topics: Elisa Bertino (Purdue University, USA) for her talk on “Privacy in the Era of Big Data, Machine Learning, IoT, and 5G”, Amit Sheth (University of South Carolina, USA) for his talk on the third wave of AI, and Torben Bach Pedersen (Aalborg University, Denmark) for his talk on “Extreme-Scale Model-Based Time Series Management with ModelarDB”.

In addition, we had a panel discussion on “Big Minds Sharing their Vision on the Future of AI” led by Bernhard Moser (SCCH, Austria), with Battista Biggio (University of Cagliari, Italy), Claudia Diaz (Katholieke Universiteit Leuven,

Belgium), Heiko Paulheim (University of Mannheim, Germany), and Olga Saukh (Complexity Science Hub, Austria).

As is the tradition of DEXA, all accepted papers were published in “Lecture Notes in Computer Science” (LNCS) and made available by Springer. Authors of selected papers presented at the conference will be invited to submit substantially extended versions of their conference papers for publication in special issues of international journals. The submitted extended versions will undergo a further review process.

The 32nd edition of DEXA featured six international workshops – three established ones and three brand-new ones – covering a variety of specific topics:

- The 12th International Workshop on Biological Knowledge Discovery from Data (BIOKDD 2021)
- The 5th International Workshop on Cyber-Security and Functional Safety in Cyber-Physical Systems (IWCFS 2021)
- The 3rd International Workshop on Machine Learning and Knowledge Graphs (MLKgraphs 2021)
- The 1st International Workshop on Artificial Intelligence for Clean, Affordable, and Reliable Energy Supply (AI-CARES 2021)
- The 1st International Workshop on Time Ordered Data (ProTime2021)
- The 1st International Workshop on AI System Engineering: Math, Modelling, and Software (AISys2021)

The success of the conference is due to the continuous and generous support of its participants and their relentless efforts. Our sincere thanks go to the dedicated authors, renowned Program Committee members, session chairs, organizing and steering committee members, and student volunteers who worked tirelessly to ensure the continuity and high quality of DEXA 2021.

We would also like to express our thanks to all institutions actively supporting this event, namely:

- Institute of Telekooperation, Johannes Kepler University Linz (JKU), Austria
- Software Competence Center Hagenberg (SCCH), Austria
- Web Applications Society (@WAS)

We hope you have enjoyed the conference! We are looking forward to seeing you again next year.

September 2021

Christine Strauss

# Organization

## Program Committee Chair

Christine Strauss

University of Vienna, Austria

## Steering Committee

Gabriele Kotsis

Johannes Kepler University Linz, Austria

A Min Tjoa

Vienna University of Technology, Austria

Robert Wille

Software Competence Center Hagenberg, Austria

Bernhard Moser

Software Competence Center Hagenberg, Austria

Ismail Khalil

Johannes Kepler University Linz, Austria

## Program Committee

Susan Ariel Aaronson

George Washington University, USA

Javier Nieves Acedo

Azterlan, Spain

Sonali Agarwal

IIIT, India

Hamid Aghajan

Ghent University, Belgium

Hans Akkermans

Vrije Universiteit Amsterdam, The Netherlands

Riccardo Albertoni

CNR-IMATI, Italy

Idir Amine Amarouche

USTHB, Algeria

Rachid Anane

Coventry University, UK

Mustafa Atay

Winston-Salem State University, USA

Sören Auer

Leibniz Universität Hannover, Germany

Juan Carlos Augusto

Middlessex University London, UK

Monica Barratt

RMIT University, Australia

Ladjel Bellatreche

LIAS, ENSMA, France

Nadia Bennani

LIRIS, INSA de Lyon, France

Karim Benouaret

Université Claude Bernard Lyon 1, France

Djamal Benslimane

Université de Lyon, France

Morad Benyoucef

University of Ottawa, Canada

Mikael Berndtsson

University of Skövde, Sweden

Catherine Berrut

LIG, Université Joseph Fourier, France

Vasudha Bhatnagar

University of Delhi, India

Didier Bigo

King's College London, UK

Steven Bird

Charles Darwin University, Australia

Ankur Singh Bist

KIET Ghaziabad, India

Joseph Bonneau

New York University, USA

Johan Bos

University of Groningen, The Netherlands

Athman Bouguettaya

University of Sydney, Australia

Olivier Bousquet

Google Brain, Zurich, Switzerland

Omar Boussai	ERIC Laboratory, France
Kevin Bowyer	University of Notre Dame, USA
Stephane Bressan	National University of Singapore, Singapore
Marcel Broersma	University of Groningen, The Netherlands
Axel Bruns	Queensland University of Technology in Brisbane, Australia
Jean Burgess	Queensland University of Technology in Brisbane, Australia
Maria Chiara Carrozza	Scuola Superiore Sant'Anna, Pisa, Italy
Lemuria Carter	University of New South Wales, Australia
Antonio Casilli	Télécom Paris, France
Pablo Castells	Universidad Autónoma de Madrid, Spain
Carlos Castillo	Universitat Pompeu Fabra, Spain
Barbara Catania	Università degli Studi di Genova, Italy
Sharma Chakravarthy	University of Texas at Arlington, USA
Max Chevalier	IRIT, France
Chen-Fu Chien	National Tsing Hua University, Taiwan
Ruzanna Chitchyan	University of Bristol, UK
Soon Ae Chun	City University of New York, USA
David Clark	MIT Computer Science and Artificial Intelligence Lab, USA
Mark Coeckelbergh	University of Vienna, Austria
Diane Cook	Washington State University, USA
Alfredo Cuzzocrea	University of Calabria, Italy
Debora Dahl	Conversational Technologies, USA
Boyd Danah	Harvard University, USA
Jérôme Darmont	Université de Lyon, France
Trevor Darrell	University of California, Berkeley, USA
Soumyava Das	Teradata Labs, USA
Robert Davison	City University of Hong Kong, Hong Kong
Emiliano De Cristofaro	University College London, UK
Ronald Deibert	University of Toronto, Canada
Vincenzo Deufemia	University of Salerno, Italy
Roberto Di Pietro	Hamad Bin Khalifa University, Qatar
Juliette Dibie-Barthélemy	INRAE, France
Dejing Dou	University of Oregon, USA
Karen Douglas	University of Kent, UK
Brian Earp	Yale School of Medicine, USA
Johann Eder	University of Klagenfurt, Austria
Nicole Ellison	University of Michigan, USA
Suzanne Embury	University of Manchester, UK
Markus Endres	University of Passau, Germany
Sergio Escalera	University of Barcelona, Spain
Charles Ess	University of Oslo, Norway
Kevin Esterling	University of California, Riverside, USA
James Evans	University of Chicago, USA



Noura Faci	Université de Lyon, France
Hany Farid	Berkeley School of Information, USA
Bettina Fazzinga	ICAR-CNR, Rende, Italy
Stefano Ferilli	Universita' di Bari, Italy
Miriam Fernandez	Open University, UK
Flavio Ferrarotti	Software Competence Centre Hagenberg, Austria
Mariel Finucane	Mathematica Policy Research, Cambridge, USA
Seth Flaxman	Imperial College London, UK
Luciano Floridi	University of Oxford, UK
Vladimir Fomichov	National Research University Higher School of Economics, Russia
Flavius FrasinCAR	Erasmus University Rotterdam, The Netherlands
Bernhard Freudenthaler	Software Competence Center Hagenberg, Austria
Andrea Fumagalli	Universita di Pavia, Italy
Steven Furnell	Plymouth University, UK
Aryya Gangopadhyay	University of Maryland Baltimore County, USA
David Garcia	Complexity Science Hub Vienna, Austria
Jorge Lloret Gazo	University of Zaragoza, Spain
David Geary	University of Missouri, USA
Claudio Gennaro	ISTI-CNR Pisa, Italy
George Gerard	Singapore Management University, Singapore
Manolis Gergatsoulis	Ionian University, Greece
Carlo Ghezzi	Politecnico di Milano, Italy
Javad Ghofrani	HTW Dresden University of Applied Sciences, Germany
Dmitry Goldgof	University of South Florida, USA
Don Gotterbarn	Access East Tennessee State University, USA
Vikram Goyal	IIIT-Delhi, India
Carmine Gravino	University of Salerno, Italy
Sven Groppe	University of Lübeck, Germany
William Grosky	University of Michigan, USA
Francesco Guerra	Università di Modena e Reggio Emilia, Italy
Giovanna Guerrini	University of Genova, Italy
Allel Hadjali	LIAS, ENSMA, France
Abdelkader Hameurlain	IRIT, Paul Sabatier University, France
Ibrahim Hamidah	Universiti Putra Malaysia, Malaysia
Takahiro Hara	Osaka University, Japan
Lynda Hardman	The Centrum Wiskunde and Informatica, The Netherlands
Eszter Hargittai	University of Zurich, Switzerland
Sven Hartmann	Clausthal University of Technology, Germany
Manfred Hauswirth	The Fraunhofer Institute for Open Communication Systems FOKUS, Germany
Eva Heiskanen	University of Helsinki, Finland
Julio Hernandez-Castro	University of Kent, UK
Antonio Hidalgo	Universidad Politécnica de Madrid, Spain

Magdalena Hurtado	Arizona State University, USA
Ionut Iacob	Georgia Southern University, USA
Sergio Ilarri	University of Zaragoza, Spain
Abdessamad Imine	Loria, France
Yasunori Ishihara	Nanzan University, Japan
Ivan Izonin	Lviv Polytechnic National University, Ukraine
Peiquan Jin	University of Science and Technology of China, China
Deborah Johnson	University of Virginia, USA
Anne Kao	Boeing, USA
Dimitris Karagiannis	University of Vienna, Austria
Stefan Katzenbeisser	TU Darmstadt, Germany
Anne Kayem	Hasso Plattner Institute, University of Potsdam, Germany
Deanna Kemp	University of Queensland, Australia
Faisal Khan	University of Calgary, Canada
Ewan Klein	University of Edinburgh, UK
Carsten Kleiner	University of Applied Science and Arts Hannover, Germany
Peter Knees	Vienna University of Technology, Austria
Henning Koehler	Massey University, New Zealand
Michal Kratky	VSB-Technical University of Ostrava, Czech Republic
Petr Kremen	Czech Technical University in Prague, Czech Republic
David Kreps	Stanford University, USA
Agnes Kukulska-Hulme	Open University, UK
Tahu Kukutai	University of Waikato, New Zealand
Josef Küng	Johannes Kepler University Linz, Austria
Nhien-An Le Khac	University College Dublin, Ireland
Lenka Lhotska	Czech Technical University in Prague, Czech Republic
Wenxin Liang	Chongqing University of Posts and Telecommunications, China
Chuan-Ming Liu	National Taipei University of Technology, Taiwan
Oscar Pastor Lopez	Universitat Politècnica de València, Spain
Hui Ma	Victoria University of Wellington, New Zealand
Qiang Ma	Kyoto University, Japan
Zakaria Maamar	Zayed University, UAE
Sanjay Madria	Missouri University of Science and Technology, USA
Elio Masciari	Federico II University, Italy
Brahim Medjahed	University of Michigan, USA
Jun Miyazaki	Tokyo Institute of Technology, Japan
Lars Moench	University of Hagen, Germany
Riad Mokadem	Paul Sabatier University, France
Anirban Mondal	University of Tokyo, Japan
Yang-Sae Moon	Kangwon National University, South Korea
Franck Morvan	IRIT, Paul Sabatier University, France
Cedric du Mouza	CNAM, France
Francesc Munoz-Escoi	Universitat Politècnica de València, Spain

Ismael Navas-Delgado	University of Malaga, Spain
Wilfred Ng	Hong Kong University of Science and Technology, Hong Kong
Marcin Paprzycki	Systems Research Institute, Polish Academy of Sciences, Poland
Dhaval Patel	IBM, USA
Clara Pizzuti	CNR-ICAR, Italy
Elaheh Pourabbas	CNR-ICAR, Italy
Uday Kiran Rage	University of Tokyo, Japan
Rodolfo Resende	Federal University of Minas Gerais, Brazil
Claudia Roncancio	Grenoble Alps University, France
Viera Rozinajova	Slovak University of Technology in Bratislava, Slovakia
Massimo Ruffolo	ICAR-CNR, Italy
Shelly Sachdeva	National Institute of Technology Delhi, India
Marinette Savonnet	University of Burgundy, France
Florence Sedes	IRIT, Paul Sabatier University, France
Nazha Selmaoui	University of New Caledonia, New Caledonia
Michael Sheng	Macquarie University, Australia
Patrick Siarry	Université de Paris 12, France
Tarique Siddiqui	Microsoft Research Lab, Redmond, USA
Gheorghe Cosmin Silaghi	Babes-Bolyai University, Romania
Hala Skaf-Molli	University of Nantes, LS2N, France
Srinivasa Srinath	IIITB, India
Bala Srinivasan	Monash University, Australia
Olivier Teste	IRIT, France
Stephanie Teufel	University of Fribourg, Switzerland
Jukka Teuhola	University of Turku, Finland
Jean-Marc Thevenin	IRIT, Université Toulouse I, France
A Min Tjoa	Vienna University of Technology, Austria
Vicenc Torra	University Skövde, Sweden
Traian Marius Truta	Northern Kentucky University, USA
Lucia Vaira	University of Salento, Italy
Ismini Vasileiou	De Montfort University, UK
Krishnamurthy Vidyasankar	Memorial University, Canada
Marco Vieira	University of Coimbra, Portugal
Piotr Wisniewski	Nicolaus Copernicus University, Poland
Ming Hour Yang	Chung Yuan Chritian University, Taiwan
Haruo Yokota	Tokyo Institute of Technology, Japan
Qiang Zhu	University of Michigan, USA
Yan Zhu	Southwest Jiaotong University, China
Ester Zumpano	University of Calabria, Italy

## External Reviewers

Tooba Aamir  
Amani Abusafia  
Abdulwahab Aljubairy  
Mohammed Bahutair  
Andrea Baraldi  
Nabila Berkani  
Francesco Del Buono  
Loredana Caruccio  
Olivier De Casanove  
Dipankar Chaki  
Rachid Chelouah  
Stefano Cirillo  
Labbe Cyril  
Matthew Damigos  
Jonathan Debure  
Abir Farouzi  
Sheik Mohammad Mostakim Fattah  
Angelo Ferrando  
Lukas Fischer  
Arnaud Flori  
Jorge Galicia  
María del Carmen Rodríguez Hernández  
Akm Tauhidul Islam  
Eleftherios Kalogeros  
Julius Köpke  
Bogdan Kostov  
Cyril Labbe  
Chuan-Chi Lai  
Hieu Hanh Le  
Xuhong Li

Ji Liu  
Jin Lu  
Qiuhaio Lu  
Jia-Ning Luo  
Jorge Martinez-Gil  
Ahcene Menasria  
Niccolo Meneghetti  
Quoc Hung Ngo  
Daria Novoseltseva  
Matteo Paganelli  
Louise Parkin  
Gang Qian  
Subhash Sagar  
Nadouri Sana  
Chayma Sellami  
Mohamed Sellami  
Vladimir A. Shekhovtsov  
Tao Shi  
Hannes Sochor  
Manel Souibgui  
Sofia Stamou  
Carlos Telleria-Orriols  
Daniele Traversaro  
Oscar Urrea  
Francesco Visalli  
Shuang Wang  
Yi-Hung Wu  
Fa Yao Yin  
Feng Yu  
Eric Zhang

## Organizers



## **Abstracts of Keynote Talks**

# Privacy in the Era of Big Data, Machine Learning, IoT, and 5G

Elisa Bertino

Samuel Conte Professor of Computer Science, Cyber2SLab, Director,  
CS Department, Purdue University, West Lafayette, Indiana, USA

**Abstract.** Technological advances, such as IoT devices, cyber-physical systems, smart mobile devices, data analytics, social networks, and increased communication capabilities are making possible to capture and to quickly process and analyze huge amounts of data from which to extract information critical for many critical tasks, such as healthcare and cyber security. In the area of cyber security, such tasks include user authentication, access control, anomaly detection, user monitoring, and protection from insider threat. By analyzing and integrating data collected on the Internet and the Web one can identify connections and relationships among individuals that may in turn help with homeland protection. By collecting and mining data concerning user travels, contacts and disease outbreaks one can predict disease spreading across geographical areas. And those are just a few examples. The use of data for those tasks raises however major privacy concerns. Collected data, even if anonymized by removing identifiers such as names or social security numbers, when linked with other data may lead to re-identify the individuals to which specific data items are related to. Also, as organizations, such as governmental agencies, often need to collaborate on security tasks, data sets are exchanged across different organizations, resulting in these data sets being available to many different parties. Privacy breaches may occur at different layers and components in our interconnected systems. In this talk, I first present an interesting privacy attack that exploits paging occasion in 5G cellular networks and possible defenses. Such attack shows that achieving privacy is challenging and there is no unique technique that one can use; rather one must combine different techniques depending also on the intended use of data. Examples of these techniques and their applications are presented. Finally, I discuss the notion of data transparency – critical when dealing with user sensitive data, and elaborate on the different dimensions of data transparency.

# Don't Handicap AI without Explicit Knowledge

Amit Sheth

University of South Carolina, USA

**Abstract.** Knowledge representation as expert system rules or using frames and variety of logics, played a key role in capturing explicit knowledge during the hay days of AI in the past century. Such knowledge, aligned with planning and reasoning are part of what we refer to as Symbolic AI. The resurgent AI of this century in the form of Statistical AI has benefitted from massive data and computing. On some tasks, deep learning methods have even exceeded human performance levels. This gave the false sense that data alone is enough, and explicit knowledge is not needed. But as we start chasing machine intelligence that is comparable with human intelligence, there is an increasing realization that we cannot do without explicit knowledge. Neuroscience (role of long-term memory, strong interactions between different specialized regions of data on tasks such as multimodal sensing), cognitive science (bottom brain versus top brain, perception versus cognition), brain-inspired computing, behavioral economics (system 1 versus system 2), and other disciplines point to need for furthering AI to neuro-symbolic AI (i.e., hybrid of Statistical AI and Symbolic AI, also referred to as the third wave of AI). As we make this progress, the role of explicit knowledge becomes more evident. I will specifically look at our endeavor to support human-like intelligence, our desire for AI systems to interact with humans naturally, and our need to explain the path and reasons for AI systems' workings. Nevertheless, the variety of knowledge needed to support understanding and intelligence is varied and complex. Using the example of progressing from NLP to NLU, I will demonstrate the dimensions of explicit knowledge, which may include, linguistic, language syntax, common sense, general (world model), specialized (e.g., geographic), and domain-specific (e.g., mental health) knowledge. I will also argue that despite this complexity, such knowledge can be scalability created and maintained (even dynamically or continually). Finally, I will describe our work on knowledge-infused learning as an example strategy for fusing statistical and symbolic AI in a variety of ways.



# Extreme-Scale Model-Based Time Series Management with ModelarDB

Torben Bach Pedersen

Aalborg University, Denmark

**Abstract.** To monitor critical industrial devices such as wind turbines, high quality sensors sampled at a high frequency are increasingly used. Current technology does not handle these extreme-scale time series well, so only simple aggregates are traditionally stored, removing outliers and fluctuations that could indicate problems. As a remedy, we present a model-based approach for managing extreme-scale time series that approximates the time series values using mathematical functions (models) and stores only model coefficients rather than data values. Compression is done both for individual time series and for correlated groups of time series. The keynote will present concepts, techniques, and algorithms from model-based time series management and our implementation of these in the open source Time Series Management System (TSMS) ModelarDB. Furthermore, it will present our experimental evaluation of ModelarDB on extreme-scale real-world time series, which shows that compared to widely used Big Data formats, ModelarDB provides up to 14x faster ingestion due to high compression, 113x better compression due to its adaptability, 573x faster aggregation by using models, and close to linear scale-out scalability.

# **Big Minds Sharing their Vision on the Future of AI (Panel)**

## **Panelists**

Battista Biggio, University of Cagliari, Italy  
Claudia Diaz, Katholieke Universiteit Leuven, Belgium  
Heiko Paulheim, University Mannheim, Germany  
Olga Saukh, Complexity Science Hub, Austria

## **Moderator**

Bernhard Moser, Software Competence Center Hagenberg and Austrian Society for Artificial Intelligence, Austria

**Abstract.** While we are currently mainly talking about narrow AI systems, in the future, neural networks will increasingly be combined with graph-based and symbolic-logical approaches (3rd wave of AI).

How will this technological trend affect the key issues of security such as integrity protection or privacy protection, and environmental impact? In this context, in this interactive panel discussion, technology experts will discuss current and envisioned challenges to AI from the research perspective of their respective fields.

# Contents – Part I

## Big Data

Reference Architecture for Running Large Scale Data Integration Experiments . . . . .	3
<i>Michał Bodziony and Robert Wrembel</i>	
Subgroup Discovery with Consecutive Erosion on Discontinuous Intervals. . .	10
<i>Reynald Eugenie and Erick Stattner</i>	
Fast SQL/Row Pattern Recognition Query Processing Using Parallel Primitives on GPUs. . . . .	22
<i>Tsubasa Ohara, Qiong Chang, and Jun Miyazaki</i>	
Scalable Tabular Metadata Location and Classification in Large-Scale Structured Datasets . . . . .	35
<i>Kazi Islam and Michael Gubanov</i>	
Unified and View-Specific Multiple Kernel K-Means Clustering. . . . .	51
<i>Yujing Zhang, Siwei Wang, and En Zhu</i>	

## Data Analysis and Data Modeling

Augmented Lineage: Traceability of Data Analysis Including Complex UDFs. . . . .	65
<i>Masaya Yamada, Hiroyuki Kitagawa, Toshiyuki Amagasa, and Akiyoshi Matono</i>	
Neural Ordinary Differential Equations for the Regression of Macroeconomics Data Under the Green Solow Model . . . . .	78
<i>Zi-Yu Khoo, Kang Hao Lee, Zhibo Huang, and Stéphane Bressan</i>	
A Quantum-Inspired Neural Network Model for Predictive BPaaS Management. . . . .	91
<i>Ameni Hedhli, Haithem Mezni, and Lamjed Ben Said</i>	
Predicting Psychiatric Diseases Using AutoAI: A Performance Analysis Based on Health Insurance Billing Data. . . . .	104
<i>Markus Bertl, Peeter Ross, and Dirk Draheim</i>	
Improving Billboard Advertising Revenue Using Transactional Modeling and Pattern Mining . . . . .	112
<i>P. Revanth Rathan, P. Krishna Reddy, and Anirban Mondal</i>	

Sarcasm Detection for Japanese Text Using BERT and Emoji . . . . .	119
<i>Yoshio Okimoto, Kosuke Suwa, Jianwei Zhang, and Lin Li</i>	
Sigmalaw PBSA - A Deep Learning Model for Aspect-Based Sentiment Analysis for the Legal Domain . . . . .	125
<i>Isanka Rajapaksha, Chanika Ruchini Mudalige, Dilini Karunaratna, Nisansa de Silva, Amal Shehan Perera, and Gathika Ratnayaka</i>	
BERT-Based Sentiment Analysis: A Software Engineering Perspective . . . . .	138
<i>Himanshu Batra, Narinder Singh Pun, Sanjay Kumar Sonbhadra, and Sonali Agarwal</i>	
A Stochastic Block Model Based Approach to Detect Outliers in Networks . . . . .	149
<i>Fabrizio Angiulli, Fabio Fassetti, and Cristina Serrao</i>	
Medical-Based Text Classification Using FastText Features and CNN-LSTM Model. . . . .	155
<i>Mohamed Walid Zeghdaoui, Omar Boussaid, Fadila Bentayeb, and Frederik Joly</i>	
<b>Data Mining</b>	
Diversified Pattern Mining on Large Graphs. . . . .	171
<i>Xin Wang, Liang Tang, Yong Liu, Huayi Zhan, and Xuanzhe Feng</i>	
EHUCM: An Efficient Algorithm for Mining High Utility Co-location Patterns from Spatial Datasets with Feature-specific Utilities. . . . .	185
<i>Yinqiao Li, Lizhen Wang, Peizhong Yang, and Junyi Li</i>	
BERT-Based Multi-Task Learning for Aspect-Based Opinion Mining . . . . .	192
<i>Manil Patel and C. I. Ezeife</i>	
GPU-Accelerated Vertex Orbit Counting for 5-Vertex Subgraphs . . . . .	205
<i>Shuya Suganami and Toshiyuki Amagasa</i>	
<b>Databases and Data Management</b>	
Efficient Discovery of Partial Periodic-Frequent Patterns in Temporal Databases. . . . .	221
<i>So Nakamura, R. Uday Kiran, P. Likhitha, P. Ravikumar, Yutaka Watanobe, Minh Son Dao, Koji Zettsu, and Masashi Toyoda</i>	
Database Framework for Supporting Retention Policies . . . . .	228
<i>Nick Scope, Alexander Rasin, James Wagner, Ben Lenard, and Karen Heart</i>	

Internal Data Imputation in Data Warehouse Dimensions . . . . .	237
<i>Yuzhao Yang, Fatma Abdelhédi, Jérôme Darmont, Franck Ravat, and Olivier Teste</i>	
Purging Data from Backups by Encryption. . . . .	245
<i>Nick Scope, Alexander Rasin, James Wagner, Ben Lenard, and Karen Heart</i>	
<b>Information Retrieval</b>	
Improving Quality of Ensemble Technique for Categorical Data Clustering Using Granule Computing . . . . .	261
<i>Rahmah Brnawy and Nematollaah Shiri</i>	
Online Optimized Product Quantization for Dynamic Database Using SVD-Updating. . . . .	273
<i>Kota Yukawa and Toshiyuki Amagasa</i>	
Querying Collections of Tree-Structured Records in the Presence of Within-Record Referential Constraints . . . . .	285
<i>Foto N. Afrati and Matthew Damigos</i>	
Dealing with Plethoric Answers of SPARQL Queries . . . . .	292
<i>Louise Parkin, Brice Chardin, Stéphane Jean, Allel Hadjali, and Mickaël Baron</i>	
<b>Prediction and Decision Support</b>	
Feature Selection and Software Defect Prediction by Different Ensemble Classifiers . . . . .	307
<i>Natalya Shakhovska and Vitaliy Yakovyna</i>	
Traffic Flow Prediction Through the Fusion of Spatial-Temporal Data and Points of Interest. . . . .	314
<i>Wanzhi Xiao, Li Kuang, and Ying An</i>	
Predicting Student Performance in Experiential Education . . . . .	328
<i>Lejia Lin, Leonard Wee Liat Tan, Nicole Hui Lin Kan, Ooi Kiang Tan, Chun Chau Sze, and Wilson Wen Bin Goh</i>	
Log-Based Anomaly Detection with Multi-Head Scaled Dot-Product Attention Mechanism. . . . .	335
<i>Qingfeng Du, Liang Zhao, Jincheng Xu, Yongqi Han, and Shuangli Zhang</i>	

Enhancing Scan Matching Algorithms via Genetic Programming  
for Supporting Big Moving Objects Tracking and Analysis  
in Emerging Environments . . . . . 348  
*Alfredo Cuzzocrea, Kristijan Lenac, and Enzo Mumolo*

A Stacking Approach for Cross-Domain Argument Identification . . . . . 361  
*Alaa Alhamzeh, Mohamed Bouhaouel, Előd Egyed-Zsigmond,  
Jelena Mitrović, Lionel Brunie, and Harald Kosch*

Property Analysis of Stay Points for POI Recommendation . . . . . 374  
*Junjie Sun, Yuta Matsushima, and Qiang Ma*

Beacon Technology for Retailers - Tracking Consumer Behavior Inside  
Brick-and-Mortar-Stores. . . . . 380  
*Alexander Voelz, Andreas Mladenow, and Christine Strauss*

**Author Index . . . . . 391**

## Contents – Part II

### Authenticity, Privacy, Security and Trust

Less is More: Feature Choosing under Privacy-Preservation for Efficient Web Spam Detection. . . . .	3
<i>Jia-Qing Wang, Yan Zhu, Huan He, and Chun-Ping Li</i>	
Construction of Differentially Private Summaries Over Fully Homomorphic Encryption . . . . .	9
<i>Shojiro Ushiyama, Tsubasa Takahashi, Masashi Kudo, and Hayato Yamana</i>	
SafecareOnto: A Cyber-Physical Security Ontology for Healthcare Systems . .	22
<i>Fatma-Zohra Hannou, Faten Atigui, Nadira Lammari, and Samira Si-said Cherfi</i>	
Repurpose Image Identification for Fake News Detection. . . . .	35
<i>Steven Jia He Lee, Tangqing Li, Wynne Hsu, and Mong Li Lee</i>	

### Data and Information Processing

An Urgency-Aware and Revenue-Based Itemset Placement Framework for Retail Stores . . . . .	51
<i>Raghav Mittal, Anirban Mondal, Parul Chaudhary, and P. Krishna Reddy</i>	
NV-QALSH: An NVM-Optimized Implementation of Query-Aware Locality-Sensitive Hashing . . . . .	58
<i>Zhili Yao, Jiaqiao Zhang, and Jianlin Feng</i>	
NCRedis: An NVM-Optimized Redis with Memory Caching . . . . .	70
<i>Jiaqiao Zhang, Zhili Yao, and Jianlin Feng</i>	
A Highly Modular Architecture for Canned Pattern Selection Problem. . . . .	77
<i>Marinos Tzanikos, Maria Krommyda, and Verena Kantere</i>	
AutoEncoder for Neuroimage . . . . .	84
<i>Mingli Zhang, Fan Zhang, Jianxin Zhang, Ahmad Chaddad, Fenghua Guo, Wenbin Zhang, Ji Zhang, and Alan Evans</i>	

## Knowledge Discovery

Towards New Model for Handling Inconsistency Issues in DL-Lite Knowledge Bases . . . . .	93
<i>Ghassen Hamdi and Mohamed Nazih Omri</i>	
ContextWalk: Embedding Networks with Context Information Extracted from News Articles . . . . .	100
<i>Chaoran Chen, Mirco Schönfeld, and Jürgen Pfeffer</i>	
FIP-SHA - Finding Individual Profiles Through SHared Accounts. . . . .	115
<i>Carolina Nery, Renata Galante, and Weverton Cordeiro</i>	
A Tag-Based Transformer Community Question Answering Learning-to-Rank Model in the Home Improvement Domain . . . . .	127
<i>Macedo Maia, Siegfried Handschuh, and Markus Endres</i>	
An Autonomous Crowdsourcing System . . . . .	139
<i>Yu Suzuki</i>	

## Machine Learning

The Effect of IoT Data Completeness and Correctness on Explainable Machine Learning Models . . . . .	151
<i>Shelernaz Azimi and Claus Pahl</i>	
Analysis of Behavioral Facilitation Tweets for Large-Scale Natural Disasters Dataset Using Machine Learning . . . . .	161
<i>Yu Suzuki, Yoshiki Yoneda, and Akiyo Nadamoto</i>	
Using Cross Lingual Learning for Detecting Hate Speech in Portuguese . . . .	170
<i>Anderson Almeida Firmino, Cláudio Souza de Baptista, and Anselmo Cardoso de Paiva</i>	
MMEnsemble: Imbalanced Classification Framework Using Metric Learning and Multi-sampling Ratio Ensemble. . . . .	176
<i>Takahiro Komamizu</i>	
Evaluate the Contribution of Multiple Participants in Federated Learning . . . .	189
<i>Zhaoyang You, Xinya Wu, Kexuan Chen, Xinyi Liu, and Chao Wu</i>	
DFL-Net: Effective Object Detection via Distinguishable Feature Learning. . .	195
<i>Jia Xie, Shouhong Wan, and Peiquan Jin</i>	
Transfer Learning for Larger, Broader, and Deeper Neural-Network Quantum States. . . . .	207
<i>Remmy Zen and Stéphane Bressan</i>	



LGTM: A Fast and Accurate kNN Search Algorithm in High-Dimensional Spaces. . . . .	220
<i>Yusuke Arai, Daichi Amagata, Sumio Fujita, and Takahiro Hara</i>	
TSX-Means: An Optimal K Search Approach for Time Series Clustering. . . .	232
<i>Jannai Tokotoko, Nazha Selmaoui-Folcher, Rodrigue Govan, and Hugues Lemonnier</i>	
A Globally Optimal Label Selection Method via Genetic Algorithm for Multi-label Classification . . . . .	239
<i>Tianqi Ji, Jun Li, and Jianhua Xu</i>	
<b>Semantic Web and Ontologies</b>	
Discovering HOI Semantics from Massive Image Data . . . . .	251
<i>Mingguang Zheng, Shouhong Wan, and Peiquan Jin</i>	
Fuzzy Ontology-Based Possibilistic Approach for Document Indexing Using Semantic Concept Relations . . . . .	264
<i>Kabil Boukhari and Mohamed Nazih Omri</i>	
Multi-Objective Recommendations and Promotions at TOTAL . . . . .	270
<i>Idir Benouaret, Mohamed Bouadi, and Sihem Amer-Yahia</i>	
An Effective Algorithm for Classification of Text with Weak Sequential Relationships . . . . .	283
<i>Qiqiang Xu, Ji Zhang, Ting Yu, Wenbin Zhang, Mingli Zhang, Yonglong Luo, Fulong Chen, and Zhen Liu</i>	
PatRIS: Patent Ranking Inventive Solutions . . . . .	295
<i>Xin Ni, Ahmed Samet, Hicham Chibane, and Denis Cavallucci</i>	
<b>Temporal, Spatial, and High Dimensional Databases</b>	
Shared-Memory Parallel Hash-Based Stream Join in Continuous Data Streams . . . . .	313
<i>Peyman Behzadnia</i>	
Event Related Data Collection from Microblog Streams . . . . .	319
<i>Manoj K. Agarwal, Animesh Baranawal, Yogesh Simmhan, and Manish Gupta</i>	
GACE: Graph-Attention-Network-Based Cardinality Estimator . . . . .	332
<i>Daobing Zhu, Dongsheng He, Shuhuan Fan, Jianming Liao, and Mengshu Hou</i>	

A Two-Phase Approach for Enumeration of Maximal  $(\Delta, \gamma)$ -Cliques  
of a Temporal Network . . . . . 346  
*Suman Banerjee and Bithika Pal*

**Author Index** . . . . . 359