

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/8637>

Abdelkader Hameurlain ·
A Min Tjoa (Eds.)

Transactions on Large-Scale Data- and Knowledge- Centered Systems XLVIII

Special Issue In Memory of
Univ. Prof. Dr. Roland Wagner

Editors-in-Chief
Abdelkader Hameurlain
IRIT, Paul Sabatier University
Toulouse, France

A Min Tjoa
IFS, Vienna University of Technology
Vienna, Austria

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISSN 1869-1994 ISSN 2510-4942 (electronic)
Transactions on Large-Scale Data- and Knowledge-Centered Systems
ISBN 978-3-662-63518-6 ISBN 978-3-662-63519-3 (eBook)
<https://doi.org/10.1007/978-3-662-63519-3>

© Springer-Verlag GmbH Germany, part of Springer Nature 2021, corrected publication 2021
Chapter “pygrametl: A Powerful Programming Framework for Easy Creation and Testing of ETL Flows” is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). For further details see license information in the chapter.

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-Verlag GmbH, DE
part of Springer Nature
The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

Preface

This special issue of our Transactions on Large-Scale Data- and Knowledge-Centered Systems (TLDKS) is dedicated to Roland Wagner, who passed away on January 3, 2020. His death was a tremendous loss to everyone who knew him and to the entire information, database, and knowledge systems community.

Roland Wagner (born in 1952) studied computer science at the University of Linz, where he also received his doctorate in 1979. In 1987 he was appointed professor for information systems at the Johannes Kepler University Linz (JKU). In 1990 he founded the Institute for Applied Oriented Knowledge Processing at JKU. In the same year, the idea to organize the DEXA conference on Database and Expert System Applications was born and realized with a few colleagues.

Roland was very excited and enthusiastic when, together with Mukesh Mohania, we planned an additional conference on Data Warehousing and Knowledge Discovery (DAWAK) to complement our DEXA conference, to bring these two communities together. Due to the success of DAWAK, which has established itself as one of the most important international scientific events on Big Data Analytics and Knowledge Discovery, it was again Roland who, together with Josef Küng and ourselves, initiated the idea of publishing the most relevant articles in this important area as a Lecture Notes in Computer Science (LNCS) journal. The idea of TLDKS was born. In 2009, the inaugural LNCS journal Transactions on Large-Scale Data and Knowledge-Centered System was published by Springer.

Roland Wagner was a very special person and scientist. He was literally driven by the spirit to make the world a better place through scientific work in his field. All of his actions were derived from this spirit. The creation and organization of conferences, the establishment of excellent centers, the cooperation with developing countries, and, last but not least, his commitment to people with special needs testify to this spirit and this driving force. With his commitment and enthusiasm, he triggered a large number of projects, initiatives, and innovations that are very well recognized in Austria and internationally. He received the Humanity Medal of Upper Austria (2003), the First Class Medal of the Prague Czech Technical University (2002), and the Upper Austrian Silver Decoration of Honor (2011). The Roland Wagner Award, founded in his honor, is presented every two years by the Austrian Computer Society at the International Conference on Computers Helping People with Special Needs (under the auspices of UNESCO) for achievements by scientists or organizations that have groundbreaking effects on the lives of people with disabilities.

Roland Wagner was a wonderful colleague who wholeheartedly supported our community. He inspired many of us immensely, and we can best honor him through our work as researchers in developing contributions that will ultimately lead to applications for a world where “no one should be left behind”. Roland is bitterly missed, but his legacy will be kept alive through the many initiatives he established, the

many collaborations he instigated, and the many students and researchers he mentored throughout the world.

This special issue includes eight invited papers addressing a range of relevant and hot topics: Distributed Database Systems, NewSQL, Scalable Transaction Management, Caches, Strong Consistency, Data Warehouse, ETL, Reinforcement Learning, Stochastic Approximation, Multi-Agent Systems, Ontology, Requirements Engineering, Model-Driven Development, Organizational Modelling, Digital Government, New Institutional Economics, and Data Governance.

The paper “*Roland Wagner - Scientist with a Big Heart*”, which appears in the front matter, synthesizes in a remarkable way the multi-faceted career and the trajectory of Roland Wagner: his main scientific contributions, the creation of the Institute for Applied Oriented Knowledge Processing at JKU, the foundation, with his friend A Min Tjoa, of the International Conference on Database and Expert Systems Applications (DEXA), the launch of this international journal TLDKS, and the establishment of bridges between several public and private institutes. These links between research and development highlight his commitment, his valuable contributions, and his interest in improving economic and societal aspects through information technology.

In the first research paper, “*Distributed Database Systems: The Case for NewSQL*”, the authors make the case for NewSQL, which is the latest technology in the big data management landscape. They describe in detail the main characteristics of NewSQL database systems, where the main objective is to combine the scalability and availability of NoSQL with the consistency and usability of classical SQL database systems. The authors illustrate the introduced principles with the Spanner and LeanXcale systems.

In the second research paper, “*Boosting OLTP Performance Using Write-Back Client-Side Caches*”, the authors present a comprehensive evaluation of client-side caches to enhance the performance of MySQL for online transaction processing (OLTP) workloads. They focus on the TPC-C benchmark and the write-back policy of the client-side cache. The authors extend the cache with a transaction processing storage manager (Berkeley DB) to minimize its amount of required memory, quantifying its impact on TPC-C’s *tpm-C*.

The third scientific contribution is focused on “*pygrametl: A Powerful Programming Framework for Easy Creation and Testing of ETL Flows*”. Here, the authors present an overview of the latest version of pygrametl, an open source Python-based framework for ETL programmers. They describe how pygrametl offers a novel approach to ETL programming by providing a framework that abstracts over the data warehouse tables while still allowing the user to use the full power of Python.

In the fourth paper, entitled “*A Data Warehouse of Wi-Fi Sessions for Contact Tracing and Outbreak Investigation*”, the authors seek to leverage the availability of common existing digital infrastructure, such as the increasingly ubiquitous Wi-Fi networks, that can be readily activated to assist in large-scale contact tracing. Then, they describe and discuss the design, implementation, and deployment of a data warehouse of Wi-Fi sessions for contact tracing and disease outbreak investigation. Finally, the authors present the case where the data warehouse of Wi-Fi sessions is experimentally deployed at full scale on a large local university campus in Singapore.

The fifth paper addresses “*Convergence Proof for Actor-Critic Methods Applied to PPO and RUDDER*”. The authors prove under commonly used assumptions the convergence of actor-critic reinforcement learning algorithms. Their framework shows convergence of the well-known Proximal Policy Optimization (PPO) algorithm and of the recently introduced Return Decomposition for Delayed Rewards (RUDDER) algorithm.

In the sixth paper, entitled “*Revival of MAS Technologies in Industry*”, the authors introduce a new MAS Platform – the Cluster 4.0 Integration Platform. It utilizes (1) semantics to explicitly describe products, production, and resources for their easier integration and exploitation and (2) OPC-UA to connect software agents to physical machines.

In the seventh paper, entitled “*From Strategy to Code: Achieving Strategic Alignment in Software Development Projects Through Conceptual Modelling*”, the authors propose S2C, a strategy-to-code methodological approach to integrate organizational, business process, and information system modelling levels to support strategic alignment in software development. They discuss how their model-driven approach not only supports strategic alignment, but fosters the elicitation of business process performance measurement requirements.

In the last paper of this special issue, entitled “*On State-Level Architecture of Digital Government Ecosystems: From ICT-Driven to Data-Centric*”, the authors systematically approach the state-level architecture of digital government ecosystems. They establish the notion of data governance architecture, which links data assets with accountable organizations. Finally, the authors describe how the proposed framework perfectly fits the current discussion on moving from ICT-driven to data-centric digital government.

We strongly hope that this special issue, dedicated to the memory of our friend and colleague Roland Wagner, presents a broad spectrum of interesting developments relating to data management systems.

We would like to sincerely thank the authors who spontaneously accepted our invitation to contribute to this special issue and without taking into account their strong constraints and their commitments elsewhere.

Our special thanks also go to Gabriela Wagner, who has always perfectly managed all the volumes of the TLDKS journal since its creation and launch in March 2009, for her efforts and availability.

April 2021

Abdelkader Hameurlain
A Min Tjoa

Organization

Editors-in-Chief

Abdelkader Hameurlain
A Min Tjoa

Paul Sabatier University, France
Technical University of Vienna, Austria

Editorial Board

Reza Akbarinia
Dagmar Auer
Djamal Benslimane
Stéphane Bressan
Mirel Cosulschi
Dirk Draheim
Johann Eder
Anastasios Gounaris
Theo Härder
Sergio Ilarri
Petar Jovanovic
Aida Kamišalić Latifć
Dieter Kranzlmüller
Philippe Lamarre
Lenka Lhotská
Vladimir Marik
Jorge Martinez Gil
Franck Morvan
Torben Bach Pedersen
Günther Pernul
Soror Sahri
Shaoyi Yin
Feng “George” Yu

Inria, France
FAW, Austria
Lyon 1 University, France
National University of Singapore, Singapore
University of Craiova, Romania
University of Innsbruck, Austria
Alpen Adria University Klagenfurt, Austria
Aristotle University of Thessaloniki, Greece
Technical University of Kaiserslautern, Germany
University of Zaragoza, Spain
Universitat Politècnica de Catalunya, Spain
University of Maribor, Slovenia
Ludwig-Maximilians-Universität München, Germany
INSA Lyon, France
Technical University of Prague, Czech Republic
Technical University of Prague, Czech Republic
Software Competence Center Hagenberg, Austria
Paul Sabatier University, France
Aalborg University, Denmark
University of Regensburg, Germany
Paris Descartes University, France
Paul Sabatier University, France
Youngstown State University, USA

Message from Bruno Buchberger: Roland Wagner - Scientist with a Big Heart

Bruno Buchberger

Research Institute for Symbolic Computation, Johannes Kepler University, Linz,
Austria

`bruno.buchberger@jku.at`

In 1974, I had just joined the Johannes Kepler University (JKU) as a young professor for computer mathematics and gave my first course on “Introduction to Theoretical Computer Science”. I will never forget two students who attended this course: A Min Tjoa and Roland Wagner. They behaved so differently from most of the other students. They would stay in the institute, sometimes just in the seminar room, until late in the evening and on weekends. They worked on their own ideas and very early started to write papers. They were excited about science and research and, by intrinsic motivation, without any external persuasion, apparently understood the fascination of university life and an academic career.

Their topics of interest were different from mine and, for some years, I lost track of their careers until, in 1989, I had moved my Research Institute for Symbolic Computation (RISC) to the renovated Castle of Hagenberg – about 20 km off the JKU main campus - with the intention to build up what I called a “Softwarepark”, i.e. a place where research and teaching institutions would grow in close interaction with start-up companies and established companies. For a quick growth, I wanted to convince a couple of my JKU colleagues to open up branches of their JKU institutes in the frame of the new Softwarepark Hagenberg.

This turned out to be not so easy because, at that time, only very few professors, in addition to doing research, wanted to work with companies on applied projects and neither wanted to think about supporting start-up companies and creating IT jobs. In this situation, Roland, who was an associate professor in the meantime, phoned me and said “super, I want to join you with my group”. And this word “super” is my main and dear memory about Roland. In the coming years, I heard “super” many times from him. Whatever new idea or plan I had to expand the growing Softwarepark, Roland said “super” and I knew I could rely on him to get it on track.

At the beginning, I had nothing to offer to him than a first small grant from the Upper Austrian Government and a room in the castle, which later, when I had raised private money for erecting more and more buildings in the Softwarepark, became the castle restaurant.

We shared rooms and we shared the pioneering spirit of just working for research and, at the same time, for very practical company projects. We were proud to show my RISC and his institute FAW (Research Institute for Application-oriented Knowledge Processing) to visitors, politicians, companies, students, and potential co-workers and, quickly, the Softwarepark started to grow. More and more companies joined and also

(very few) other JKU institutes, notably Peter Klement's FLLL (Fuzzy Logic Laboratory Linz) and Gustav Pomberger's Software Engineering Group. As a team of four JKU institutes, we then founded a couple of other, grant-based, research institutions as, for example, the new FH Hagenberg (University of Applied Sciences, Faculty of Informatics), the Software Competence Center Hagenberg, the High School for Communication, and others.

And every time some new ideas and initiatives came to our mind, Roland said "super" and embarked on them with full energy and optimism.

Also, Roland invented and realized quite some projects, institutions, and initiatives on his own, which showed not only his strength as an internationally renowned and recognized researcher and scientist but also his interest and engagement for the societal and economic implications of IT and his big heart, in particular, for those who need support and special attention. Thus, together with his Ph.D. student and later professor Klaus Miesenberger, he initiated a project for helping visually impaired students, which grew into the "Institut Integriert Studieren". He also was a driving force behind the International Conference on Computers Helping People with Special Needs. Together with the Mayor of Hagenberg, Rudolf Fischlehner, we also made it possible to host a company in the frame of the Softwarepark that employs exclusively people with special needs, who feel completely integrated into the vibrating life of a technology park.

The days of joint work are over but Roland's spirit will stay forever with us. When I visited him in the hospital and just felt a little disappointed about everyday problems in the Softwarepark he reminded me that we had a "super" time together and that our spirit should always strive for a "super" positive attitude. Roland, I feel very grateful for what you did for the community and for me personally and I will always remember you as the "super" man!

Contents

Distributed Database Systems: The Case for NewSQL	1
<i>Patrick Valduriez, Ricardo Jimenez-Peris, and M. Tamer Özsu</i>	
Boosting OLTP Performance Using Write-Back Client-Side Caches	16
<i>Shahram Ghandeharizadeh, Haoyu Huang, and Hieu Nguyen</i>	
pygrametl: A Powerful Programming Framework for Easy Creation and Testing of ETL Flows	45
<i>Søren Kejser Jensen, Christian Thomsen, Torben Bach Pedersen, and Ove Andersen</i>	
A Data Warehouse of Wi-Fi Sessions for Contact Tracing and Outbreak Investigation	85
<i>Guilherme Augusto Zagatti, See-Kiong Ng, and Stéphane Bressan</i>	
Convergence Proof for Actor-Critic Methods Applied to PPO and RUDDER	105
<i>Markus Holzleitner, Lukas Gruber, José Arjona-Medina, Johannes Brandstetter, and Sepp Hochreiter</i>	
Revival of MAS Technologies in Industry	131
<i>Petr Kadera, Václav Jirkovský, and Vladimír Mařík</i>	
From Strategy to Code: Achieving Strategic Alignment in Software Development Projects Through Conceptual Modelling	145
<i>Oscar Pastor, Rene Noel, and Ignacio Panach</i>	
On State-Level Architecture of Digital Government Ecosystems: From ICT-Driven to Data-Centric	165
<i>Dirk Draheim, Robert Krimmer, and Tanel Tammet</i>	
Correction to: pygrametl: A Powerful Programming Framework for Easy Creation and Testing of ETL Flows	C1
<i>Søren Kejser Jensen, Christian Thomsen, Torben Bach Pedersen, and Ove Andersen</i>	
Author Index	197