Advances in Database Technology
— EDBT 2021

24th International Conference on Extending Database Technology
Nicosia, Cyprus, March 23–26, 2021
Proceedings

Editors
Yannis Velegrakis
Demetris Zeinalipour
Panos K. Chrysanthis
Francesco Guerra
Foreword by the PC Chair

The International Conference on Extending Database Technology (EDBT) is an established forum for researchers and practitioners alike to disseminate knowledge and research results related to data management. This year, the 24th edition of EDBT, scheduled to take place between March 23rd and March 26, 2021 in Nicosia, Cyprus, has instead been held entirely online, due to the circumstances and the travel restrictions imposed by the 2020/2021 pandemic. It has been jointly organized with the International Conference on Database Theory (ICDT).

The organizing committee solicited contributions in many different areas, including, but not limited to, Data Preparation, Data Privacy, Database Engines, Distributed Data Systems, Graph Management, Reasoning over Data, Machine Learning and AI in Databases, Novel Database Architectures, Semi-structured Data Management, and User Interfaces for Data. Novel in this year’s edition is the solicitation of contributions in the area of Applied Database Systems for Data Science. The goal is to give the opportunity to researchers from different areas dealing with interesting data management challenges in the context of Data Science, to disseminate them to the data management community, and at the same time to give database researchers working on interesting data science scenarios to publish their works.

The main Research track, the Industrial & Application track, as well as the Demonstration track have remained as in previous years, while the Short paper track has been slightly extended to accommodate papers of 6 pages, to allow more technical content.

With some rare exceptions, research papers were reviewed by 4 reviewers. Discussions and decisions were taken under the coordination of a senior PC member, with expertise in the respective area of the paper under review. The reviewing phase involved two cycles: resulting in a number of papers being revised and having their quality improved significantly.

Consistent with the tradition, the program included two keynotes, one on Data Profiling by Felix Naumann (Hasso Plattner Institute, Germany), and one on Knowledge Management by Katja Hose (Aalborg University, Denmark). These keynotes were complemented by the two additional keynotes of the co-organized ICDT. Last but not least, the program had the traditional tutorial track, featuring four tutorials on topics related to knowledge graphs, blockchains, text-to-SQL, and time series management. The overall program was accompanied by 6 Workshops.

The Program committee reviewed 108 full research papers, of which 27 were accepted. For short papers, 34 papers out of 113 were selected. The Industry and Application track received 22 submissions of which it selected 11 for publication, while the demo track received 24 works of which 15 were accepted.

Given that this year the conference took place online, it was a good opportunity to exploit at the maximum the opportunities that digital technologies can offer. Furthermore, it was our intention to make sure that the paper contributions are disseminated to a broader audience, especially outside the conference participants. For this reason, the authors were asked to provide a pre-recorded presentation of 10 min that will remain in the proceedings, a 30 sec pitch video that advertises the results of the contribution and a graphics ad.

In order to recognize significant contributions and give credits to the authors, the technical program committee awarded the Best Paper award to one of the research papers and the Best Demonstration award to one of the demos. Furthermore, following the EDBT tradition, it awarded the Test-of-Time award to a paper from the EDBT 2011 proceedings that has been deemed to have the greatest impact among those of that year. A novelty in this year’s EDBT edition is the additional recognition of the Best Short Paper.

The realization of EDBT 2021 is a result of a collaborative effort of the different chairs and program committee members. Congratulations are in place to the senior PC members for guiding the discussions in such a professional and timely manner, ensuring the selection of the best quality papers. Special thanks for the excellent collaboration and quality of work go to the Industrial and Application Chair Eric Simon, the Demonstration Chair Sihem Amer-Yahia, the tutorial chairs Stefan Manegold and Wang-Chiew Tan, and the Workshop Chair Evaggelia Pitoura. A great deal of credits go to the Proceedings Chair Francesco Guerra for all the extra work he has put in organizing the material and making sure that all the proceedings information is in place. My appreciation goes also to the members of the paper award committees for the effort they put in evaluating the candidate papers and providing the decisions in a timely manner. I find it impossible to describe the passion, professionalism, consistency and collaborative attitude the general chairs Demetris Zeinalipour and Panos K. Chrysanthis have demonstrated. It was great working with them. I would also like to thank all the program committee members since due to them the high quality program became possible, the award committees, as well as Angela Bonifati and Marc H. Scholl from the EDBT Board for the their numerous advices and support. Last but not least, I would like to thank all the authors for the works they submitted to the conference, the keynote speakers, the tutorial presenters, and the demonstration presenters.

Yannis Velegrakis, EDBT 2021 PC Chair
**Message from the General Chairs**

The 24th edition of the International Conference on Extending Database Technology (EDBT) was held between the 23rd and 26th of March 2021, despite the COVID-19 worldwide pandemic. Continuing its longstanding tradition as a premier data management research forum taking place in a European location, EDBT 2021, jointly organized with the International Conference on Database Theory (ICDT), was held virtually in Nicosia, the capital of the beautiful island of Cyprus.

At the onset, our aim has been to offer as close to an in-person, face-to-face experience as possible, maximizing the dissemination of knowledge and sharing of research results among the EDBT/ICDT participants at no health risk to them, and together discover the Cypriot culture and hospitality. EDBT/ICDT 2021 was initially planned to become the first hybrid EDBT/ICDT conference, combining physical presentations with an extended online audience, although this plan later evolved into an online format due to the ongoing situation with COVID-19. To this end, we have been committed to offering the best possible online experience to attendees by capitalizing and expanding on the success of earlier conferences.

Particularly, we decided on a number of novelties compared to previous conferences. Firstly, given that an optimal online experience cannot be achieved with pre-recorded presentations, but only with live presentations and direct interactions between the speakers and the audience, we decided in coordination with the Program Chairs on the online synchronous delivery of all events, including all research presentations, tutorials and live demonstrations. The next challenge was to select the appropriate conference management platform. Our personal experience as participants in several online conferences during the past year, and our evaluation of a number of popular web socializing platforms, led us to the decision that a custom conference attendance experience can only be delivered by an in-house platform that relies on a reliable teleconferencing channel, an intuitive interface, and a low learning curve.

To this end, the Data Management Systems Laboratory at the University of Cyprus, under the leadership of Demetris Zeinalipour developed and hosted a novel web-based platform named VGATE (Virtual Gate) to online conferences. VGATE allowed the Organizing Committee to collaborate over a Web-based Sheet interface (like Google Sheets) to plan and manage the organization. VGATE provided numerous helpful building-blocks in the online organization, namely, zoom license management and user status integration, integration of proceedings and multimedia content, industry booths, live sessions, social rooms, and for the first time an online helpdesk supported by Easy-Conferences Ltd. Special thanks to Paschalis Mpeis (University of Cyprus, Cyprus), Soteris Constantinoiu (University of Cyprus, Cyprus) and Constantinos Costa (University of Pittsburgh, USA) for their support and code contributions to the project under an extremely tight schedule. Special thanks also to Zoom Video Communications, Inc., for facilitating and sponsoring EDBT/ICDT 2021.

Novel this year is the participation of EDBT/ICDT to the newly founded Diversity and Inclusion (D&I) initiative of the Data Management community. EDBT/ICDT (alongside SIGMOD, VLDB, SoCC, and ICDE) celebrates the diversity in our community and welcomes everyone regardless of age, sex, gender identity, race, ethnicity, socioeconomic background, country of origin, religion, sexual orientation, physical ability, education, work experience, etc. To introduce this initiative, Panos K. Chrysanthis, the EDBT/ICDT D&I Chair, together with Sihem Amer-Yahia (CNRS, University Grenoble Alpes, France), a D&I Core Member, organized a panel as part of the Reception on DAY1 of the conference.

With D&I and broad participation in mind, the conference program was structured in a way that is convenient to at least half of the world at any given time. Specifically, the program was split into morning sessions (CET), which were convenient for EU-ASIA participants, on DAY1 and DAY3, and the afternoon sessions (CET), which were convenient for EU-America participants on DAY2 and DAY4. The Workshops were scheduled on DAY1 in the morning or afternoon based on the geographical location of their presenters. Social events were split along the same lines, presenting the host country, Cyprus, through music and videos on culture, geography, food, leisure, and other enjoyable aspects of Cyprus. Virtual corridor and hallway discussions were supported by a number of unmoderated and dynamic social rooms on VGATE at all times.

This year we also introduced some additional novelties, namely: (i) We introduced a special ceremony during Reception on DAY2, titled “Pandemic Greetings from the Pioneers and the Next Data Management Challenge”, with greetings by Philip A. Bernstein (Microsoft Research, WA, USA), Laura M. Haas (University of Massachusetts – Amherst, MA, USA), Yannis Ioannidis (University of Athens, Greece) and Jeffrey D. Ullman (Stanford, CA, USA); (ii) we introduced the concept of a sponsored industry talk during Dinner on DAY3, with title “Behind the Scenes of Snowflake’s new Search Optimization Service”, by Ismail Oukid (Snowflake, Germany) and Stefan Richter (Snowflake, Germany); (iii) we introduced a dedicated “Demos in Action: Meet the Authors!” session during Dinner on DAY3, which allowed Demo presenters to individually showcase their demos to participants in their private presentation space through...
VGATE. Finally, we also continued to support the Climate Change discussion with a session on DAY4, led by Antoine Amarilli (Télécom Paris, France), with guest Benjamin Pierce (University of Pennsylvania, USA).

In all above endeavors, we had the support and encouragement of our incredible Organization Committee and the EDBT Executive Board, in particular, the President of EDBT Executive Board Angela Bonifati (Lyon 1 University, France), as well as the Chair of the ICDT Council Wim Martens (University of Bayreuth, Germany).

We are grateful to the entire Technical Program Committees, which, under the excellent leadership of the EDBT Program Chair Yannis Velegrakis (University of Trento, Italy and Utrecht University, Netherlands) and the ICDT Program Chair Ke Yi (Hong Kong University of Science and Technology, Hong Kong), brought forward an exciting technical program with a rich production of technical content (proceedings, videos, pitches, ads, etc.). Our gratitude also extends personally to the EDBT Demonstration Chair Sihem Amer-Yahia (CNRS, University Grenoble Alpes, France), the EDBT Workshop Chair Evaggelia Pitoura (University of Ioannina, Greece), the Climate Chair Antoine Amarilli (Télécom Paris, France), the EDBT Applied Database Systems for Data Science Vice-Chair Paul Groth (University of Amsterdam, Netherlands), the EDBT Industrial/Application Chair Eric Simon (SAP, France), the Tutorial Chairs Stefan Manegold (CWI, Netherlands) and Wang-Chiew Tan (Megagon Labs, USA), for the excellent collaboration, and exchange of ideas and insightful discussions. We would also like to highlight EDBT Demonstration Chair Sihem Amer-Yahia’s success in putting together EDBT’s first all-women Demonstration Track Committee. We would also like to thank the organizers of the six, co-located workshops DOLAP, BigVis, BMDA, DARLI-AP, SIMPLIFY and PIE+Q for enriching the scope of the technical program.

Several people contributed to the successful organization of the EDBT/ICDT 2021 conference. Special thanks to the following valued collaborators for their passion in organizing a memorable event: the Sponsorship Chair Divyakant Agrawal (University of California-Santa Barbara, USA), the Publicity Chair Herodotos Herodotou (Cyprus University of Technology, Cyprus), the Finance Chair George Pallis (University of Cyprus, Cyprus), the EDBT Proceedings Chair Francesco Guerra (University of Modena and Reggio Emilia, Italy), the ICDT Proceedings Chair Zhewei Wei (Renmin University, China), the Workshops Proceedings Chair Constantinos Costa (University of Pittsburgh, USA) and the Website support by Nicolas Kantzilaris (Easy Conferences, Cyprus).

Our sincere gratitude to our platinum sponsor, Snowflake, and our bronze sponsors, Oracle and Zoom, as well as the contact persons behind the support, namely Martin Hentschel (Snowflake), Ann Brisson (Oracle) and Alberto Colautti (Zoom).

Organizing EDBT 2021 at the University of Cyprus was Prof. George Samaras†’ passion. After his unexpected passing two years ago, his friends and colleagues at the University of Cyprus volunteered to make his passion a reality. We are grateful to the EDBT Executive Board for giving us this opportunity and trusting us to organize EDBT 2021 alongside ICDT 2021 in Cyprus, as proposed by George in 2018. We dedicate the EDBT/ICDT 2021 in honor of the memory of Prof. George Samaras† (1959–2018).

We hope you enjoyed EDBT/ICDT 2021’s exciting technical program and your virtual visit to Cyprus!

Demetris Zeinalipour, University of Cyprus, Cyprus
Panos K. Chrysanthis, University of Cyprus, Cyprus and University of Pittsburgh, USA

EDBT/ICDT 2021 General Co-Chairs
# Program Committee Members

## Research Program Committee Chair
Yannis Velegrakis, U Trento, Italy & Utrecht U, The Netherlands

## Senior Program Committee Members

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periklis Andritsos</td>
<td>U Toronto, Canada</td>
</tr>
<tr>
<td>Nikos Bikakis</td>
<td>Athena Res. Ctr., Greece</td>
</tr>
<tr>
<td>Elena Ferrari</td>
<td>U Insurbria, Italy</td>
</tr>
<tr>
<td>Avigdor Gal</td>
<td>Technion – Israel IT, Israel</td>
</tr>
<tr>
<td>Lukasz Golaz</td>
<td>U Waterloo, Canada</td>
</tr>
<tr>
<td>Paul Groth</td>
<td>U Amsterdam, Netherlands</td>
</tr>
<tr>
<td>Sergio Greco</td>
<td>U Calabria, Italy</td>
</tr>
<tr>
<td>Christian S. Jensen</td>
<td>Aalborg U, Denmark</td>
</tr>
<tr>
<td>Laks V.S. Lakshmanan</td>
<td>U British Columbia, Canada</td>
</tr>
<tr>
<td>Qiong Luo</td>
<td>HKUST, China</td>
</tr>
<tr>
<td>Raymond Ng</td>
<td>U British Columbia, Canada</td>
</tr>
<tr>
<td>Tamer Özsu</td>
<td>U Waterloo, Canada</td>
</tr>
<tr>
<td>Pinar Tozun</td>
<td>IT U Copenhagen, Denmark</td>
</tr>
</tbody>
</table>

## Program Committee Members

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karl Aberer</td>
<td>EPFL, Switzerland</td>
</tr>
<tr>
<td>Ashraf Aboulnaga</td>
<td>Qatar Comp. Res. Inst., HBKU, Qatar</td>
</tr>
<tr>
<td>Bernd Amann</td>
<td>Sorbonne U, France</td>
</tr>
<tr>
<td>Nicolas Anciaux</td>
<td>INRIA, France</td>
</tr>
<tr>
<td>Walid G. Aref</td>
<td>Purdue U, USA</td>
</tr>
<tr>
<td>Akhil Arora</td>
<td>EPFL, Switzerland</td>
</tr>
<tr>
<td>Manos Athanassoulis</td>
<td>Bost U, USA</td>
</tr>
<tr>
<td>Nikolaus Augsten</td>
<td>U Salzburg, Austria</td>
</tr>
<tr>
<td>Elena Baralis</td>
<td>Politecnico di Torino, Italy</td>
</tr>
<tr>
<td>Denilson Barbosa</td>
<td>U Alberta, Canada</td>
</tr>
<tr>
<td>Ilaria Bartolini</td>
<td>U Bologna, Italy</td>
</tr>
<tr>
<td>Senjuti Basu Roy</td>
<td>New Jersey CTU, USA</td>
</tr>
<tr>
<td>Luigi Bellomarini</td>
<td>Banca d’Italia, Italy</td>
</tr>
<tr>
<td>Michael Benedikt</td>
<td>U Oxford, UK</td>
</tr>
<tr>
<td>Alexander Boehm</td>
<td>SAP SE, Germany</td>
</tr>
<tr>
<td>Luc Bouganim</td>
<td>INRIA, France</td>
</tr>
<tr>
<td>Andrea Cali</td>
<td>Birkbeck U London, UK</td>
</tr>
<tr>
<td>Marco Calautti</td>
<td>U Trento, Italy</td>
</tr>
<tr>
<td>Bogdan Cautis</td>
<td>U Paris-Sud, France</td>
</tr>
<tr>
<td>Mel Chekol</td>
<td>Utrecht U, Netherlands</td>
</tr>
<tr>
<td>Vassilis Christophides</td>
<td>ENSEA, ETIS, France</td>
</tr>
<tr>
<td>Dario Colazzo</td>
<td>U Paris Dauphine – PSL, France</td>
</tr>
<tr>
<td>Bin Cui</td>
<td>Peking U, China</td>
</tr>
<tr>
<td>Alfredo Cuzzocrea</td>
<td>ICAR-CNRS, U Calabria, Italy</td>
</tr>
<tr>
<td>Sabrina De Capitani</td>
<td>Vimercati, U Milano, Italy</td>
</tr>
<tr>
<td>Antonios Deligiannakis</td>
<td>TU Crete, Greece</td>
</tr>
<tr>
<td>Çağatay Demiralp</td>
<td>Sigma Computing, USA</td>
</tr>
<tr>
<td>Stefania Dumbrava</td>
<td>ENSIEE, France</td>
</tr>
<tr>
<td>George Fakas</td>
<td>Uppsala U, Sweden</td>
</tr>
<tr>
<td>Ju Fan</td>
<td>Renmin U China, China</td>
</tr>
<tr>
<td>Donatella Firmani</td>
<td>Roma Tre U, Italy</td>
</tr>
<tr>
<td>George Fletcher</td>
<td>Eindhoven UT, Netherlands</td>
</tr>
<tr>
<td>Daniele Foroni</td>
<td>Huawei, Germany</td>
</tr>
<tr>
<td>Johann-Christoph Freytag</td>
<td>HU Berlin, Germany</td>
</tr>
<tr>
<td>Avigdor Gal</td>
<td>Technion – Israel IT, Israel</td>
</tr>
<tr>
<td>Johann Gamper</td>
<td>Free U Bozen-Bolzano, Italy</td>
</tr>
<tr>
<td>Yunjun Gao</td>
<td>Zhejiang U, China</td>
</tr>
<tr>
<td>Rainer Gemulla</td>
<td>U Mannheim, Germany</td>
</tr>
<tr>
<td>Boris Glavic</td>
<td>Illinois IT, USA</td>
</tr>
<tr>
<td>Paolo Guagliardo</td>
<td>U Edinburgh, UK</td>
</tr>
<tr>
<td>Michael Guibonov</td>
<td>Florida State U, USA</td>
</tr>
<tr>
<td>Xi He</td>
<td>U Waterloo, Canada</td>
</tr>
<tr>
<td>Melanie Herschel</td>
<td>U Stuttgart, Germany</td>
</tr>
<tr>
<td>Jan Hidders</td>
<td>U London, UK</td>
</tr>
<tr>
<td>Katja Hose</td>
<td>Aalborg U, Denmark</td>
</tr>
<tr>
<td>Vagelis Hristidis</td>
<td>U California – Riverside, USA</td>
</tr>
<tr>
<td>Haoyu Huang</td>
<td>Google, USA</td>
</tr>
<tr>
<td>Xin Huang</td>
<td>Hong Kong Baptist U, Hong Kong SAR</td>
</tr>
<tr>
<td>Ekaterini Ioanou</td>
<td>Tilburg U, Netherlands</td>
</tr>
<tr>
<td>Zsolt István</td>
<td>ITU Copenhagen, Denmark</td>
</tr>
<tr>
<td>Panos Kalnis</td>
<td>King Abdullah UST, Saudi Arabia</td>
</tr>
<tr>
<td>Vana Kalogeraki</td>
<td>Athens U Eco. &amp; Busin., Greece</td>
</tr>
<tr>
<td>Verena Kantere</td>
<td>National TU Athens, Greece</td>
</tr>
<tr>
<td>Panagiotis Karras</td>
<td>Aarhus U, Denmark</td>
</tr>
<tr>
<td>Asterios Katsifodinos</td>
<td>TU Delft, Netherlands</td>
</tr>
<tr>
<td>Anastasios Kementsietsidis</td>
<td>Google Research, USA</td>
</tr>
<tr>
<td>Haridimos Kondylakis</td>
<td>FORTH-ICS, Greece</td>
</tr>
<tr>
<td>Nick Koudas</td>
<td>U Toronto, Canada</td>
</tr>
<tr>
<td>Georgia Koutrika</td>
<td>Athena Research Center, Greece</td>
</tr>
<tr>
<td>Alexandros Labrinidis</td>
<td>U Pittsburgh, USA</td>
</tr>
<tr>
<td>Ulf Leser</td>
<td>HU Berlin, Germany</td>
</tr>
<tr>
<td>Guoliang Li</td>
<td>Tsinghua U, China</td>
</tr>
<tr>
<td>Han Li</td>
<td>Amazon, USA</td>
</tr>
<tr>
<td>Xiang Lian</td>
<td>Kent State U, USA</td>
</tr>
<tr>
<td>Chunbin Lin</td>
<td>Amazon AWS, USA</td>
</tr>
<tr>
<td>Matteo Lissandrini</td>
<td>Aalborg U, Denmark</td>
</tr>
<tr>
<td>Eric Lo</td>
<td>Chinese U Hong Kong, Hong Kong SAR</td>
</tr>
<tr>
<td>Ping Lu</td>
<td>BeiHang U, China</td>
</tr>
<tr>
<td>Nikos Mamoulis</td>
<td>U Ioannina, Greece</td>
</tr>
<tr>
<td>Ioana Manolescu</td>
<td>INRIA &amp; Inst. Poly. Paris, France</td>
</tr>
<tr>
<td>Sebastian Michel</td>
<td>TU Kaiserslautern, Germany</td>
</tr>
<tr>
<td>Paolo Missier</td>
<td>Newcastle U, UK</td>
</tr>
</tbody>
</table>
Mohamed Mokbel, U Minnesota – Twin Cities, USA
Mirella M. Moro, U Federal Minas Gerais, Brazil
Davide Mottin, Aarhus U, Denmark
Hieu Nguyen, eBay, USA
Behrooz Omidvar-Tehrani, LIG, France
Mourad Ouzzani, Qatar Comp. Res. Inst., HBKU, Qatar
George Papadakis, U Athens, Greece
Olga Papaemmanouil, Brandeis U, USA
Odysseas Papapetrou, TU Eindhoven, Netherlands
Paolo Papotti, Eurecom, France
Alok Pareek, Striim, USA
Torben Bach Pedersen, Aalborg U, Denmark
Eric Peukert, Leipzig U, Germany
Dimitris Plexousakis, ICS-FORTH, Greece
Laura Po, U Modena & Reggio Emilia, Italy
Arnaud Prat, Sparsity Technologies, Spain
Nicoleta Preda, U Versailles, France
Abdulhakim Qahtan, Utrecht U, Netherlands
Mohammad Sadoghi, U California, Davis, USA
Carlo Sartiani, U Basilicata, Italy
Kai-Uwe Sattler, TU Ilmenau, Germany
Sebastian Schelter, U Amsterdam, Netherlands
Steffi Scherzinger, U Passau, Germany
Petra Selmer, Neo4j, UK
Juan Sequeda, data.world, USA
Lidan Shou, Zhejiang U, China
Giovanni Simonini, U Modena & Reggio Emilia, Italy
Hala Skaf-Molli, U Nantes, France
Kostas Stefanidis, Tampere U, Finland
Gabor Szarnyas, CWI, Netherlands
Letizia Tanca, Politecnico di Milano, Italy
Ernest Teniente, U Politècnica Catalunya, Spain
Arash Termehchy, Oregon State U, USA
Chao Tian, Alibaba, China
Riccardo Torlone, Roma Tre U, Italy
Farouk Toumani, Clermont Auvergne U, CNRS, France
Katerina Tzompanaki, Cergy Paris U, France
Vasilis Vassalos, Athens U Eco. & Busin., Greece
Panos Vassiliadis, U Ioannina, Greece
Yannis Velegrakis, Utrecht U, Netherlands
Jianguo Wang, Purdue U, USA
Wendy Hui Wang, Stevens IT, USA
Wolfram Wingerath, Baqend, Germany
Nikolay Yakovets, TU Eindhoven, Netherlands
Meihui Zhang, Beijing IT, China

Industrial Track Program Committee

Tyler Akidau, Snowflake, USA
Minhea Andrei, SAP, France
Angela Bonifati, U Lille, France
Jianjun Chen, Bytedance US Lab, USA
Thomas Fanghaenel, Salesforce, USA
Avrilia Floratou, Microsoft, USA
Prasanta Ghosh, Microsoft, USA
Yash Govind, Informatica LLC, USA
Laura Haas, U Mass. Amherst, USA
Zack Ives, U Pennsylvania, USA
Martin Kersten, MonetDB Solutions, Netherlands
Hariraharan Lakshmanan, Oracle, USA
Dustin Lange, Amazon Research, USA
Jyoti Leeka, Microsoft, USA
Stefan Mandl, EXASOL AG, Germany
Anisooa Nica, SAP Labs Waterloo, Canada
Berthold Reinwald, IBM Research-Almaden, USA
Alejandro Salinger, SAP SE, Germany
Dennis Shasha, New York U, USA
Mohamed Soliman, Datometry, USA
Nesime Tatbul, Intel Labs and MIT, USA
Wei Wang, Nat. U Singapore, Singapore
Xuezhi Wang, Google, USA
Yongsik Yoon, Snowflake, USA
Kai Zeng, Alibaba Group, China

Demonstration Track Program Committee

Anastasia Ailamaki, EPFL, Switzerland
Elena Baralis, Polytechnico di Torino, Italy
Senjuti Basu Roy, NJIT, USA
Angela Bonifati, Lyon U, France
Renata Borovica-Gajic, U Melbourne, Australia
Malu Castellanos, TERADATA, USA
Sarah Cohen Boulakia, U Paris Sud, France
Maria Luisa Damiani, U Milan, Italy
Anna Fariha, UMass Amherst, USA
Irina Fundulaci, FORTH, GRnet, Greece
Katja Hose, Aalborg U, Denmark
Vana Kalogeraki, Athens U Eco. & Busin., Greece
Zoi Kaoudi, TU Berlin, Germany
Georgia Koutrika, ATHENA, Greece
Ioana Manolescu, INRIA, France
Renée J. Miller, Northeastern U, USA
Silvia Mittel, U Maine, USA
Fatma Özcan, Google, USA
Nesime Tatbul, Intel Labs & MIT, USA
Pinar Tözün, ITU, Denmark
Esther Pacitti, U Montpellier (INRIA & CNRS), France
Danica Porobic, Oracle, Switzerland
Agma Traina, ICMC-USP, Brazil
Zografoula Vagena, U Paris, France and RelationalAI, USA
Xiaolan Wang, Megagon Labs, USA
Karine Zeitouni, U Versailles, France
Additional Reviewers

Yaniv Gur, IBM Research-Almaden, USA
Ahmed Al-Baghdadi, Kent State U, USA
Niranjan Rai, Kent State U, USA
Weilong Ren, Kent State U, USA
Tahereh Arabghalizi, U Pittsburgh, USA
Evangelos Karraygeorgos, U Pittsburgh, USA
Xiaoting Li, U Pittsburgh, USA
Anthony Sicilia, U Pittsburgh, USA
Prithu Banerjee, U BC, Vancouver, Canada
Saket Gurukar, Ohio State U, Ohio, USA
Garima Gaur, IIT Kanpur, India
Sainyam Galhotra, U Mass., Amherst, USA

Nikos Giatrakos, TU Crete, Greece
Angjela Davitkova, TU Kaiserslautern, Germany
Leonardo Gazzarri, U Stuttgart, Germany
Flavio Giobergia, Politecnico di Torino, Italy
Eliana Pastor, Politecnico di Torino, Italy
Uta Störl, Darmstadt U of Applied Sciences, Germany
Wolfgang Mauerer, TU of Appl. Sc. Regensburg, Germany
Vasilis Efthymiou FORTH ICS Greece
Nikos Myrtakis U Crete, Greece
Benjamin Wollmer, Baqend and U Hamburg, Germany
Andrea Rossi, U Roma Tre, Italy
Tobias Lindaaker, Neo4j, Sweden
Conference Organization

General Chairs
Demetris Zeinalipour, University of Cyprus, Cyprus
Panos K. Chrysanthis, University of Cyprus, Cyprus and University of Pittsburgh, USA

EDBT Program Chair
Yannis Velegrakis, University of Trento, Italy and Utrecht University, Netherlands

ICDT Program Chair
Ke Yi, Hong Kong University of Science and Technology, Hong Kong

EDBT Industrial/Application Chair
Eric Simon, SAP, France

EDBT Applied Database Systems for Data Science Vice-Chair
Paul Groth, University of Amsterdam, Netherlands

EDBT Demonstrations Chair
Sihem Amer-Yahia, CNRS, Univ. Grenoble Alpes, France

Tutorial Chairs
Stefan Manegold, CWI, Netherlands
Wang-Chiew Tan, Megagon Labs, USA

Workshops Chair
Evaggelia Pitoura, University of Ioannina, Greece

Climate Chair
Antoine Amarilli, Télécom Paris, France

EDBT Proceedings Chair
Francesco Guerra, University of Modena and Reggio Emilia, Italy

ICDT Proceedings Chair
Zhewei Wei, Renmin University, China

Workshops Proceedings Chair
Constantinos Costa, University of Pittsburgh, USA

Diversity and Inclusion Chair
Panos K. Chrysanthis, University of Pittsburgh, USA

Sponsorship Chair
Divyakant Agrawal, UC Santa Barbara, USA

Publicity Chair
Herodotos Herodotou, Cyprus University of Technology, Cyprus

Finance Chair
George Pallis, University of Cyprus, Cyprus

Website Chair
Kyriakos Georgiades, EasyConferences, Cyprus
Test-of-Time Award

Established in 2014, the Test-of-Time Award of the Extended Database Technology (EDBT) Conference recognizes papers presented at the EDBT Conferences that have had the most impact in terms of research, methodology, conceptual contribution, or transfer to practice over the past ten years. The 2021 Test-of-Time Award committee looked and evaluated the impact of the papers in the EDBT 2011 proceedings, and selected

**SeMiTri: a framework for semantic annotation of heterogeneous trajectories**
by Zhixian Yan, Dipanjan Chakraborty, Christine Parent, Stefano Spaccapietra, and Karl Aberer
published in the EDBT 2011 Proceedings, pp. 259–270, DOI: 10.1145/1951365.1951398,

because it is one of the earliest papers to propose a general method for enriching moving object trajectories with semantics useful for supporting location-based services, which have been and still are in high demand across several application sectors. Since its publication, SeMiTri has generated significant interest, and follow-up work on semantic processing of mobile data and trajectories.

The EDBT 2021 Test of Time Award committee was formed by Barbara Catania, University of Genova, Italy, Gautam Das, University of Texas at Arlington, USA, Beng Chin OOI, National University of Singapore, Singapore, Themis Palpanas, University of Paris, France, and Yufei Tao, Chinese University of Hong Kong, China.

The EDBT Test-of-Time award for 2021 will be presented during the EDBT/ICDT 2021 Conference in Nicosia, Cyprus, as part of the Awards session on March 24, 2021.
Best Paper Award

The Best Paper Award Committee has looked at the papers accepted in the conference and selected one that was distinguishing itself in terms of research quality, presentation, technical challenges, and novelty. The selected paper is

**DomainNet: Homograph Detection for Data Lake Disambiguation**

by Aristotelis Leventidis, Laura Di Rocco, Wolfgang Gatterbauer, Renée J. Miller and Mirek Riedewald.

DOI: 10.5441/002/edbt.2021.03

The paper presents DomainNet, a system that disambiguates values from heterogeneous datasets by creating a network representing co-occurring values and computing their graph centrality. The system is unsupervised, its accuracy outperforms the state-of-the-art, and it is accompanied by an open benchmark. The paper is of high significance: the problem is important, the proposed solution is effective, and the benchmark facilitates further research.

**Abstract:** Modern data lakes are deeply heterogeneous in the vocabulary that is used to describe data. We study a problem of disambiguation in data lakes: how can we determine if a data value occurring more than once in the lake has different meanings and is therefore a homograph? While word and entity disambiguation have been well studied in computational linguistics, data management and data science, we show that data lakes provide a new opportunity for disambiguation of data values since they represent a massive network of interconnected values. We investigate to what extent this network can be used to disambiguate values. DomainNet uses network-centrality measures on a bipartite graph whose nodes represent values and attributes to determine, without supervision, if a value is a homograph. A thorough experimental evaluation demonstrates that state-of-the-art techniques in domain discovery cannot be re-purposed to compete with our method. Specifically, using a domain discovery method to identify homographs has a precision and a recall of 38% versus 69% with our method on a synthetic benchmark. By applying a network-centrality measure to our graph representation, DomainNet achieves a good separation between homographs and data values with a unique meaning. On a real data lake our top-200 precision is 89%.

The EDBT 2021 Best Paper Award committee was formed by Avigdor Gal, Technion Israel Institute of Technology, Israel, Lucasz Golab, University of Waterloo, Canada, Christian Jensen, Aalborg University, Denmark, and Qiong Luo, HKUST, China.

The EDBT Best Paper Award for 2021 will be presented during the EDBT/ICDT 2021 Conference in Nicosia, Cyprus, on March 24, 2021.
Best Short Paper Award

The Best Short Paper Award Committee has looked at the papers accepted in the conference and selected one that was distinguishing itself in terms of research quality, presentation, technical challenges, novelty, and potential impact to the broader research community and industry. The selected paper is

**Answer Graph: Factorization Matters in Large Graphs**
by Zahid Abul-Basher, Nikolay Yakovets, Parke Godfrey, Stanley Clark, and Mark Chignell.
DOI: 10.5441/002/edbt.2021.56

The paper proposes a two-step method for answering SPARQL conjunctive queries. The first step constructs an answer graph consisting only of node-edge-node triples that are answers to the query (a factorized answer set). In the second step, this answer graph is used to compute the actual embeddings for the query. The authors present detailed implementations for generating answer graphs and computing the final embedding, particularly in the case of cyclic conjunctive queries. Some first results from an experimental evaluation are also provided.

The EDBT 2021 Best Short Paper Award committee was formed by Manos Athanasoulis, Boston University, USA, Johann Gamper, Free University of Bozen-Bolzano, Italy, Ioana Manolescu, INRIA, France, Letizia Tanca, University of Milan, Italy, and Yannis Kotidis, Athens University of Economics and Business, Greece

The EDBT Best Short Paper Award for 2021 will be presented during the EDBT/ICDT 2021 Conference in Nicosia, Cyprus.
Best Demonstration Award

The Best Demonstration Award Committee has reviewed the video recordings of the 15 demos accepted at EDBT/ICDT and selected the most engaging demonstration that serves as an example of a structured and well illustrated demonstration and showcases an end-to-end system on a state-of-the-art topic that promotes data management beyond its boundaries. The selected demo is

**Conversational OLAP in Action**
by Matteo Francia, Enrico Gallinucci, and Matteo Golfarelli
(DISI – University of Bologna)
DOI: 10.5441/002/edbt.2021.74

For demonstrating COOL, a tool supporting natural language COnversational OLap sessions. COOL interprets and translates a natural language dialogue into an OLAP session that starts with a GPSJ query. The demonstration is engaging and showcases the usability of COOL and its capabilities in assisting query formulation and ambiguity resolution.

The EDBT 2021 Best Demonstration Award committee was formed by Sihem Amer-Yahia, CNRS Univ. Grenoble Alpes, France, Elena Baralis, Politecnico di Torino, Italy, Maria Luisa Damiani, University of Milan, Italy, Anna Fariha, UMass Amherst, USA, Irini Fundulaki, FORTH, GRnet, Greece, Zoi Kaoudi, TU Berlin, Germany, Georgia Koutrika, ATHENA, Greece, Esther Pacitti, University of Montpellier (Inria&CNRS), France, and Agma Traina, ICMC-USP, Brazil.

The EDBT Best Demonstration Award for 2021 will be presented during the EDBT/ICDT 2021 Conference in Nicosia, Cyprus, on March 24, 2021.
Table of Contents

Foreword by the PC Chair ................................................................. i
Message from the General Chairs ..................................................... ii
Program Committee Members ........................................................ iv
Conference Organization ................................................................. vii
Test-of-Time Award ...................................................................... viii
Best Paper Award ......................................................................... ix
Best Short Paper Award ............................................................... x
Best Demonstration Award ............................................................ xi
Table of Contents ........................................................................... xii

Research Papers

Exchanging Data under Policy Views
   Angela Bonifati, Ugo Comignani, Efthymia Tsamoura ........................ 1

DomainNet: Homograph Detection for Data Lake Disambiguation
   Aristotelis Leventidis, Laura Di Rocco, Wolfgang Gatterbauer, Renée J. Miller, Mirek Riedewald ............................... 13

GPU-INSCY: A GPU-Parallel Algorithm and Tree Structure for Efficient Density-based Subspace Clustering
   Jakob Rødsgaard Jørgensen, Katrine Scheel, Ira Assent ................................. 25

JIT happens: Transactional Graph Processing in Persistent Memory meets Just-In-Time Compilation
   Muhammad Attahir Jibril, Alexander Baumstark, Philipp Götze, Kai-Uwe Sattler .................................................. 37

Fixing Wikipedia Interlinks Using Revision History Patterns
   Tova Milo, Slava Novgorodov, Kathy Razmadze ................................ 49

Automating Data Quality Validation for Dynamic Data Ingestion
   Sergey Redyuk, Zoi Kaoudi, Volker Markl, Sebastian Schelter ................................. 61

Provenance-Based Algorithms for Rich Queries over Graph Databases
   Yann Ramusat, Silvia Maniu, Pierre Senellart ........................................ 73

Sequence detection in event log files
   Ioannis Mavroudopoulos, Theodoros Toliopoulos, Christos Bellas, Andreas Kosmatopoulos, Anastasios Gounaris ................................. 85

A Comparative Evaluation of Anomaly Explanation Algorithms
   Nikolaos Myrtakis, Vassilis Christophides, Eric Simon ........................................ 97

Scaling Density-Based Clustering to Large Collections of Sets
   Daniel Kocher, Nikolaus Augsten, Willi Mann ........................................ 109

Assess Queries for Interactive Analysis of Data Cubes
   Matteo Francia, Matteo Golfarelli, Patrick Marcel, Stefano Rizzi, Panos Vassiliadis ........................................ 121

SolveDB+: SQL-Based Prescriptive Analytics
   Laurynas Siksnys, Torben Bach Pedersen, Thomas Dyhre Nielsen, Davide Frazzetto ........................................ 133

Multi-Objective Influence Maximization
   Shay Gershtein, Tova Milo, Brit Youngmann ........................................ 145

Subjectivity Aware Conversational Search Services
   Yacine Gaci, Jorge Ramirez, Boualem Benatallah, Fabio Casati, Khalid Benabdeslem ........................................ 157
GeoBlocks: A Query-Cache Accelerated Data Structure for Spatial Aggregation over Polygons
Christian Winter, Andreas Kipf, Christoph Anneser, Eleni Tziriti Zacharatou, Thomas Neumann, Alfons Kemper 169

Indoor Spatial Queries: Modeling, Indexing, and Processing
Tiantian Liu, Huan Li, Hua Lu, Muhammad Aamir Cheema, Lidan Shou 181

Structure Detection inVerbose CSV Files
Lan Jiang, Gerardo Vitagliano, Felix Naumann 193

FRESQUE: A Scalable Ingestion Framework for Secure Range Query Processing on Clouds
Hoang Tran Van, Tristan Allard, Laurent d’Orazio, Amr El Abbadi 205

Cache on Track (CoT): Decentralized Elastic Caches for Cloud Environments
Victor Zakhary, Lawrence Lim, Divy Agrawal, Amr El Abbadi 217

Knowledge Graph Management on the Edge
Weiqin Xu, Olivier Curé, Philippe Calvez 229

PolyFit: Polynomial-based Indexing Approach for Fast Approximate Range Aggregate Queries
Zhe Li, Tsz Nam Chan, Man Lung Yiu, Christian Jensen 241

Shift-Table: A Low-latency Learned Index for Range Queries using Model Correction
Ali Hadian, Thomas Heinis 253

An Efficient and Secure Location-based Alert Protocol using Searchable Encryption and Huffman Codes
Sina Shaham, Gabriel Ohtinta, Cyrus Shahabi 265

Concealer: SGX-based Secure, Volume Hiding, and Verifiable Processing of Spatial Time-Series Datasets
Peevush Gupta, Sharad Mehrotra, Shantanu Sharma, Nalini Venkatasesubramanian, Guoxi Wang 277

Evaluation of Hardening Techniques for Privacy-Preserving Record Linkage
Martin Franke, Ziad Sehili, Florens Rohde, Erhard Rahm 289

Proof-of-Execution: Reaching Consensus through Fault-Tolerant Speculation
Suyash Gupta, Jelle Hellings, Sajjad Rahnama, Mohammad Sadoghi 301

Scalable Linear Algebra Programming for Big Data Analysis
Leonidas Fegaras 313

Short Papers

Automated Machine Learning for Entity Matching Tasks
Matteo Paganelli, Francesco Del Buono, Marco Pevarello, Francesco Guerra, Maurizio Vincini 325

COCOA: Correlation COefficient-Aware Data Augmentation
Mahdi Esmailoghli, Jorge Arnulfo Quiane Ruiz, Ziawasch Abedjan 331

Efficient Exploratory Clustering Analyses with Qualitative Approximations
Manuel Fritz, Dennis Tschechlov, Holger Schwarz 337

AutoML4Clust: Efficient AutoML for Clustering Analyses
Dennis Tschechlov, Manuel Fritz, Holger Schwarz 343

Feature-driven Time Series Clustering
Donato Tiano, Angela Bonifati, Raymond Ng 349

Indexed Log File: Towards Main Memory Database Instant Recovery
Arlino Magalhães, Angelo Brayner, José Maria Monteiro, Gustavo Moraes 355

HorsePower: Accelerating Database Queries for Advanced Data Analytics
Hanfeng Chen, Joseph D’silva, Laurie Hendren, Bettina Kemme 361
Robust and Memory-Efficient Database Fragment Allocation for Large and Uncertain Database Workloads
Rainer Schlosser, Stefan Halfpap .................................................. 367

TD-AC: Efficient Data Partitioning based Truth Discovery
Mouhamadou Lamine BA, Osias Noel Nicodème Finagnon Tossou .................................................. 373

Towards Automated Concept-based Decision Tree Explanations for CNNs
Radwa El Shawi, Youssef Sherif, Sherif Sakr .................................. 379

Efficient Maintenance of Distance Labelling for Incremental Updates in Large Dynamic Graphs
Muhammad Farhan, Qing Wang .................................................. 385

KISS - A fast kNN-based Importance Score for Subspaces
Anna Beer, Ekaterina Allerborn, Valentin Hartmann, Thomas Seidl .................................................. 391

SceneRec: Scene-Based Graph Neural Networks for Recommender Systems
Gang Wang, Ziyi Guo, Xiang Li, Dawei Yin, Shuai Ma .................................................. 397

Efficient Contact Similarity Query over Uncertain Trajectories
Xichen Zhang, Suprio Roy, Farzaneh Shoelhe, Rongxing Lu .................................................. 403

AdCom: Adaptive Combiner for Streaming Aggregations
Felipe Gutierrez, Kaustubh Beedkar, Abel Souza, Volker Markl .................................................. 409

Querying Top-k Dominant Traffic Flows on Large Urban Road Networks
Stella Maropaki, Paolo Sottovia, Stefano Bortoli .................................................. 415

On Supporting Scalable Active Learning-based Interactive Data Exploration with Uncertainty Estimation Index
Xiaoyu Ge, Panos Chrysanthis .................................................. 421

Efficient Discovery of Approximate Order Dependencies
Reza Karegar, Parke Godfrey, Lukasz Golab, Mehdi Kargar, Divesh Srivastava, Jaroslav Szlichta .................................................. 427

Towards Scalable Data Discovery
Javier Flores, Sergi Nadal, Oscar Romero .................................................. 433

Adaptive Multi-Model Reinforcement Learning for Online Database Tuning
Yaniv Gur, Dongsheng Yang, Frederik Stalschus, Berthold Reinwald .................................................. 439

Optimising Fairness Through Parametrised Data Sampling
Vladimiro González-Zelaya, Julián Salas, Dennis Prangle, Paolo Missier .................................................. 445

Using Landmarks for Explaining Entity Matching Models
Andrea Baraldi, Francesco Del Buono, Matteo Paganelli, Francesco Guerra .................................................. 451

Human-Interpretable Rules for Anomaly Detection in Time-Series
Ines Ben Kraiem, Faiza Ghozzi, André Péninou, Geoffrey Roman-Jimenez, Olivier Teste .................................................. 457

DBMS Performance Troubleshooting in Cloud Computing Using Transaction Clustering
Arunprasad Marathe .................................................. 463

Revisiting Multidimensional Adaptive Indexing [Experiment & Analysis]
Anders Hammershøj Jensen, Frederik Lauridsen, Fatemeh Zardbani, Stratos Idreos, Panagiotis Karras .................................................. 469

Twin Subsequence Search in Time Series
Georgios Chatzigeorgakidis, Dimitrios Skoutas, Kostas Patroumpas, Themis Palpanas, Spiros Athanasiou, Spiros Skiadopoulos .................................................. 475

Progressive Mergesort: Merging Batches of Appends into Progressive Indexes
Pedro Holanda, Stefan Manegold .................................................. 481

Multiple-Source Context-Free Path Querying in Terms of Linear Algebra
Arseniy Terekhov, Vlada Pogozhelskaya, Vadim Abzalov, Timur Zininatulin, Semyon Grigorev .................................................. 487
Answer Graph: Factorization Matters in Large Graphs

Zahid Abul-Basher, Nikolay Yakovets, Parke Godfrey, Stanley Clark, Mark Chignell

Schema Inference for Property Graphs

Hanâ Lbath, Angela Bonifati, Russ Harmer

Optimizing SPARQL Queries using Shape Statistics

Kashif Rabbani, Matteo Lissandrini, Katja Hose

Preserving Diversity in Anonymized Data

Mostafa Milani, Yu Huang, Fei Chiang

Automatic Tuning of Read-Time Tolerances for Optimized On-Demand Data-Streaming from Sensor Nodes

Julius Hülsmann, Chiao-Yun Li, Jonas Traub, Volker Markl

SOJA: A Memory-efficient Small-large Outer Join for MPI

Liang Liang, Guang Yang, Thomas Heinis, David Taniar

Industrial Papers

JENGA - A Framework to Study the Impact of Data Errors on the Predictions of Machine Learning Models

Sebastian Schelter, Tammo Rukat, Felix Biessmann

Decongestant: A Breath of Fresh Air for MongoDB Through Freshness-aware Reads

Chenhao Huang, Michael Cahill, Alan Fekete, Uwe Roehm

DLC: A New Compaction Scheme for LSM-tree with High Stability and Low Latency

Peiquan Jin, Jianchuang Li, Hai Long

Financial Data Exchange with Statistical Confidentiality: A Reasoning-based Approach

Luigi Bellomarini, Livia Blasi, Rosario Laurendi, Emanuel Sallinger

Generating Realistic Test Datasets for Duplicate Detection at Scale Using Historical Voter Data

Fabian Panse, André Düjon, Wolfram Wingerath, Benjamin Wollmer

Path Indexing in the Cypher Query Pipeline

Jochem Kuijpers, George Fletcher, Tobias Lindaaker, Nikolay Yakovets

A Deep Learning Architecture for Audience Interest Prediction of News Topic on Social Media

Ciprian-Octavian Truică, Elena-Simona APOSTOL, Teodor Ștefu, Panos Karras

AutoDBaaS: Autonomous Database as a Service for managing relational database services

Mayank Tiwary, Pritish Mishra, Shashank Mohan Jain, Kshira Sahoo

Scalable Spatio-temporal Indexing and Querying over a Document-oriented NoSQL Store

Nikolaos Koutroumanis, Christos Doulkeridis

Production Experiences from Computation Reuse at Microsoft

Alekh Jindal, Shi Qiao, Hiren Patel, Abhishek Roy, Jyoti Leeka, Brandon Haynes

WILSON: A Divide and Conquer Approach for Fast and Effective News Timeline Summarization

Yiming Liao, Shuguang Wang, Dongwon Lee

Demos

Conversational OLAP in Action

Matteo Francia, Enrico Gallinucci, Matteo Golfarelli

Smart City Data Analysis via Visualization of Correlated Attribute Patterns

Yuya Sasaki, Keizo Hori, Daiki Nishihara, Ohashi Sora, Yusuke Wakuta, Kei Harada, Makoto Onizuka, Yuki Arase, Shinji Shimojo, Kenji Doi, Hongdi He, Zhong-ren Peng
SciNeM: A Scalable Data Science Tool for Heterogeneous Network Mining
Serafeim Chatzopoulos, Thanasis Vergoulis, Panagiotis Deligiannis, Dimitrios Skoutas, Theodore Dalamagas, Christos Tryfonopoulos ................................................................. 654

IMCF: The IoT Meta-Control Firewall for Smart Buildings
Soteris Constantinou, Antonis Vasileiou, Andreas Konstantinidis, Panos Chrysanthis, Demetrios Zeinalipour-Yazti 658

BBoxDB Streams: Distributed Processing of Real-World Streams of Position Data
Jan Kristof Nidzwetzki, Ralf Hartmut Güting ................................................................. 662

Correlation graph analytics for stock time series data
Tong Liu, Paolo Coletti, Anton Dignös, Johann Gamper, Maurizio Murgia ................................................................. 666

Conquering a Panda’s weaker self - Fighting laziness with laziness
Stefan Hagedorn, Steffen Kläbe, Kai-Uwe Sattler ................................................................. 670

DocDesign 2.0: Automated Database Design for Document Stores with Multi-criteria Optimization
Moditha Hewasinghage, Sergi Nadal, Alberto Abelló ................................................................. 674

Visualizing and Exploring Big Datasets based on Semantic Community Detection
Maria Krommyda, Konstantinos Tsitseklis, Verena Kantere, Vasileios Karyotis, Symeon Papavassiliou ................. 678

Exploration and Analysis of Temporal Property Graphs
Christopher Rost, Kevin Gomez, Philip Fritzsche, Andreas Thor, Erhard Rahm ................................. 682

Coronis: Towards Integrated and Open COVID-19 Data
Giorgos Santipantakis, George Vouros, Christos Doulkeridis ................................................................. 686

Effective and Scalable Data Discovery with NextiaJD
Javier Flores, Sergi Nadal, Oscar Romero ................................................................. 690

A Tool for JSON Schema Witness Generation
Lyes Attouche, Mohamed-Amine Baazizi, Dario Colazzo, Francesco Falleni, Giorgio Ghelli, Cristiano Landi, Carlo Sartiani, Stefanie Scherzinger ................................................................. 694

covRew: a Python Toolkit for Pre-Processing Pipeline Rewriting Ensuring Coverage Constraint Satisfaction
Chiara Accinelli, Barbara Catania, Giovanna Guerrini, Simone Minisi ................................................................. 698

EasyBDI: Near Real-Time Data Analytics over Heterogeneous Data Sources
Bruno Silva, Jose Moreira, Rogério Luís Costa ................................................................. 702

Tutorials

Tutorial on the Internals of Permissioned Blockchains and on How to Build Applications with Hyperledger Fabric
Zsolt István ................................................................. 706

Deep Learning Approaches for Text-to-SQL Systems
George Katsogiannis-Meimarakis, Georgia Koutrika ................................................................. 710

Big Sequence Management: Scaling up and Out
Karima Echihabi, Kostas Zoumpatianos, Themis Palpanas ................................................................. 714