## **EDBT Conference Proceedings volume editors**

## 1. EDBT PC Chair

Yannis Velegrakis, University of Trento and Utrecht University, Netherlands

## 2. EDBT/ICDT General Chair

Demetris Zeinalipour, University of Cyprus, Cyprus

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

## 6. EDBT Proceedings Chair

Francesco Guerra, University of Modena and Reggio Emilia, Italy

## **Foreword**

A Message from the Program 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 is 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 the year before. The Short paper track has been slightly extended to accommodate papers of 6 pages, instead of 4. This brings the works accepted in that track further from the typical plenary poster works, and closer to that of full papers, giving more space to authors to provide more details of the related theory, motivations and solutions. Some of the papers submitted in the main research track that were describing significant research ideas, but deemed not yet mature to constitute a full paper, were passed to the short paper track after a second round of review.

With some rare exceptions, research papers were reviewed by 4 reviewers. Discussions and decisions were taken under the coordination of senior PC members, with expertise in the respective area of the papers under review. The reviewing phase involved two cycles. After the first phase, some papers were asked to be revised and go through a second round of review in order to be brought to the quality level required by the conference.

The program contained two keynote speakers, 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. Furthemore, 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 the significant collaborative effort of 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](https://www.cwi.nl/people/stefan-manegold) and [Wang-Chiew Tan](https://megagon.ai/team/wang-chiew-tan/), and the Workshop Chair Evaggelia Pitoura. Special credits go to the Proceedings Chair Francesco Guerra for all the extra work that 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 papers and providing the decisions in a timely manner. I also find it impossible to describe the passion, professionalism, consistency and collaborative attitude that the general chairs Demetris Zeinalipour and Panos Chrysanthis have demonstrated. It has been an honor working with them. Last but not least, we should not forget that the conference is successful due to the high quality content that has been produced by the authors, the keynote speakers, the tutorial presenters, and the demonstration presenters. And of course, the selection of such good content has become possible only with the effort of the reviewers. I would like to also thank Angela Bonifati for serving as our liaison towards the EDBT Board.

Yannis Velegrakis

EDBT 2021 Program Chair

## A Message from The General Chairs

## 

## 

## Award Committees

### Test-of-Time Award

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

Yufei Tao, Chinese University of Hong Kong, China

### Best Paper Award

**Avigdor Gal**, Technion Israel Institute of Techn., Israel

**Lucasz Golab**, University of Waterloo, Canada

**Christian Jensen**, Aalborg University , Denmark

**Qiong Luo**, HKUST, China

### Best Short Paper Award

**Manos Athanasoulis**, Boston University, USA

**Johann Gamper**, Free University of Bozen-Bolzano, Italy

**Ioana Manolescu**, INRIA, France

**Letizia Tanca**, University of Milan, Italy

**Yannis Kotidis**, Athens University of Economics and Business, Greece

## 

## Conference Organization

### General Chairs

[**Demetris Zeinalipour**](https://www2.cs.ucy.ac.cy/~dzeina), University of Cyprus, Cyprus

[**Panos K. Chrysanthis**](https://panos.cs.pitt.edu/), University of Cyprus, Cyprus and University of Pittsburgh, USA

### EDBT Program Chair

[**Yannis Velegrakis**](http://www.cs.uu.nl/staff/velegrak.html), University of Trento, Italy and Utrecht University, Netherlands

### ICDT Program Chair

[**Ke Yi**](https://www.cse.ust.hk/~yike/), Hong Kong University of Science and Technology, Hong Kong

### EDBT Industrial/Application Chair

[**Eric Simon**](https://people.sap.com/eric.simon), SAP, France

### EDBT Applied Database Systems for Data Science Vice-Chair

[**Paul Groth,**](https://www.uva.nl/en/profile/g/r/p.t.groth/p.t.groth.html) University of Amsterdam, Netherlands

### EDBT Demonstration Chair

[**Sihem Amer-Yahia**](http://lig-membres.imag.fr/amery/), CNRS, Univ. Grenoble Alpes, France

### Tutorial Chairs

[**Stefan Manegold**](https://www.cwi.nl/people/stefan-manegold), CWI, Netherlands

[**Wang-Chiew Tan**](https://megagon.ai/team/wang-chiew-tan/), Megagon Labs, USA

### Workshop Chair

[**Evaggelia Pitoura**](http://www.cs.uoi.gr/~pitoura/), University of Ioannina, Greece

### Climate Chair

[**Antoine Amarilli**](https://a3nm.net/index.html.en), Télécom Paris, France

### EDBT Proceedings Chair

[**Francesco Guerra**](http://morespace.unimore.it/francescoguerra/), University of Modena and Reggio Emilia, Italy

### ICDT Proceedings Chair

[**Zhewei Wei**](https://weizhewei.com/), Renmin University, China

### Workshops Proceedings Chair

[**Constantinos Costa**](https://people.cs.pitt.edu/~costa.c/), University of Pittsburgh, USA

### Diversity and Inclusion Chair

**Panos K. Chrysanthis**, University of Pittsburgh, USA

### Sponsorship Chair

[**Divyakant Agrawal**](https://sites.cs.ucsb.edu/~agrawal/), UC Santa Barbara, USA

### Publicity Chair

[**Herodotos Herodotou**](https://www.cut.ac.cy/faculties/fet/eecei/staff/herodotos.herodotou/?languageId=1), Cyprus University of Technology, Cyprus

### Finance Chair

[**George Pallis**](https://www.cs.ucy.ac.cy/~gpallis/), University of Cyprus, Cyprus

### Website

[**Kyriakos Georgiades**](http://easyconferences.eu/), 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 the:

**SeMiTri: a framework for semantic annotation of heterogeneous trajectories,**

by Zhixian Yan, Dipanjan Chakraborty, Christine Parent, Stefano Spaccapietra, and Karl Aberer

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 abstract of the specific paper was reading as follows:

GPS devices allow recording the movement track of the mov- ing object they are attached to. This data typically consists of a stream of spatio-temporal (x,y,t) points. For application purposes the stream is transformed into finite subsequences called trajectories. Existing knowledge extraction algorithms defined for trajectories mainly assume a specific context (e.g. vehicle movements) or analyze specific parts of a trajectory (e.g. stops), in association with data from chosen geographic sources (e.g. points-of-interest, road networks). We investi- gate a more comprehensive semantic annotation framework that allows enriching trajectories with any kind of semantic data provided by multiple 3rd party sources.

This paper presents SeMiTri - the framework that enables annotating trajectories for any kind of moving objects. Do- ing so, the application can benefit from a “semantic trajec- tory” representation of the physical movement. The frame- work and its algorithms have been designed to work on tra- jectories with varying data quality and different structures, with the objective of covering abstraction requirements of a wide range of applications. Performance of SeMiTri has been evaluated using many GPS datasets from multiple sources – including both fast moving objects (e.g. cars, trucks) and people’s trajectories (e.g. with smartphones). These two kinds of experiments are reported in this paper.

## 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 the:

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

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

*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.*

## Best Short Paper Award

The Best Short Paper award committee was formed to give credit to the specific short paper that presents some novel idea that opens many opportunities for challenging technical solutions, is well presented, and appears to have a great impact in the future. Given this, the selection committee had selected the:

## Table of Contents

Marc