The Workshops of the EDBT/ICDT 2026 Joint Conference

March 24, 2026

Message from the Proceedings Chairs:

It is our great pleasure to present on behalf of the entire conference organizing committee and the workshop organizers, the proceedings of the Workshops co-located with the 29th International Conference on Extending Database Technology (EDBT) and the 29th International Conference on Database Theory (ICDT), held on March 24, 2026 in Tampere, Finland.

The EDBT and ICDT series of conferences are prestigious forums for exchanging novel results that extend the foundations and applications of data management technologies. This year, eight exciting workshops continue the tradition of focusing on emerging topics in data management, complementing the areas covered by the main technical program (these proceedings include the first six workshops, while the last one runs its own proceedings and the Young Researcher Symposium provides a solid foundation for networking and cultivates a supportive and inclusive environment for PhD students):

- 10th International Workshop on Data Analytics solutions for Real-LIfe APplications (DARLI-AP)
- 8th International Workshop on Big Mobility Data Analytics (BMDA)
- 5th International Workshop on Data Systems Education (DataEd)
- 2nd International Workshop on Transforming Graph Data (TGD)
- 1st International Workshop on Explainable Data Science and Machine Learning for the Sciences (XAI4Science)
- 1st International Workshop on Quality in Large Language Models and Knowledge Graphs (QuaLLM-KG)
- 28th International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data (DOLAP)
- Young Researcher Symposium at EDBT 2026

We thank the workshop organizers, PC members and external reviewers for their effort in organizing these workshops, and the authors for continuing to submit their high-quality work to the EDBT/ICDT workshops, making these venues successful and intellectually stimulating.

Sincerely,
Alexander Krause, Technische Universität Dresden (Germany)
João Felipe Pimentel, Universidade Federal Fluminense (Brazil)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2026-04-09

Data Analytics solutions for Real-LIfe APplications (DARLI-AP)

Two trends are transforming technology today: the explosion of data collected by digital devices and rapid advances in data science, machine learning, and deep learning. Together, they enable new approaches across many domains, but also raise practical questions about how to turn algorithms into systems that work reliably in real settings.

DARLI-AP is a forum where academics and practitioners share methods and experience for building analytics-driven applications across their full lifecycle. The focus is on solutions that remain dependable, interpretable, and adaptable in operation, and that link models to data quality, pipelines, domain constraints, monitoring, explainability, and long-term maintenance.

Now celebrating its 10th anniversary, DARLI-AP also provides a collective perspective on how data-driven applications have evolved, from early predictive analytics and feature engineering to large-scale pipelines, deep learning, and foundation models. Across these shifts, the workshop highlights enduring open challenges: reproducible experimentation at scale, data-centric and provenance-aware practices, integration of domain knowledge and causal reasoning, robustness under distribution shift, human-in-the-loop design, and responsible deployment under constraints of ethics, regulation, sustainability, and equity.

In this tenth edition, DARLI-AP features 17 papers by 68 authors (approximately 23.5% female and 76.5% male), covering stages from data preparation to modelling, deployment, and continuous improvement. These contributions advance the state of the art while helping clarify what is already solid, what remains fragile, and which new scientific and engineering directions are needed to build more reliable, inclusive, and impactful data-driven services.

The DARLI-AP program features two keynote speeches by Prof. Ira Assent, Aarhus University, Sweden, “Reliable Explanations for Data Analytics” and Prof. Francesca Dragotto, University of Roma Tor Vergata, “Artificial tools and ‘natural’ social exclusion: a linguistic perspective”.

The organizers of DARLI-AP would like to express their heartfelt thanks to all those who contributed to the success of the tenth edition:

- The authors, for submitting their research papers to the workshop;
- The keynote speakers, Prof. Ira Assent and Prof. Francesca Dragotto, for honoring us with presentations of their recent research activities and perspectives at DARLI-AP 2026;
- The members of the Program Committee and the external reviewers, for generously dedicating their time and expertise to providing constructive and highly valuable feedback to the authors;
- The EDBT/ICDT 2026 Chairs, for their trust and their valuable support.

- Carmem Hara — Federal University of Paraná (Brazil)
- Chen Jiang — Auburn University (USA)
- Patrick Marcel — University of Orléans (France)
- Sara Migliorini — Università degli Studi di Verona (Italy)
- Simone Monaco — Politecnico di Torino (Italy)
- Santiago Negrete-Yankelevich — Universidad Autónoma Metropolitana Cuajimalpa (Mexico)
- Kjetil Nørvåg Norwegian — University of Science and Technology (Norway)
- Eliana Pastor — Politecnico di Torino (Italy)
- Elisa Quintarelli — Università di Verona (Italy)
- Simona E. Rombo — University of Palermo (Italy)
- Domenico Ursino — Polytechnic University of the Marche (Italy)
- José Luis Zechinelli Martini — Universidad de las Américas Puebla (Mexico)
- Ester Zumpano — University of Calabria (Italy)

Big Mobility Data Analytics (BMDA)

From spatial to spatio-temporal and, then, to mobility data. So, what's next? It is the rise of mobility-aware integrated Big Data analytics. The Big Mobility Data Analytics (BMDA) workshop series, initiated in 2018 with EDBT Conference, aims at bringing together experts in the field from academia, industry and research labs to discuss the lessons they have learned over the years, to demonstrate what they have achieved so far, and to plan for the future of mobility.

In its 8th edition, the BMDA workshop will foster the exchange of new ideas on multidisciplinary real-world problems, discuss proposals about innovative solutions, and identify emerging opportunities for further research in the area of big mobility data analytics, such as deep learning on mobility data, edge computing, visual analytics, etc. The workshop intends to bridge the gap between researchers and big mobility data stakeholders, including experts from critical domains, such as urban / maritime / aviation transportation, human complex networks, etc.

BMDA acknowledges the support of the following EU Horizon projects and organizations:

- EMERALDS (Extreme-scale Urban Mobility Data Analytics as a Service, EU Horizon Programme, 2023-25)

- Green.Dat.AI (Energy-efficient AI-ready Data Spaces, EU Horizon Programme, 2023-25)

Program Committee Chairs:

Anita Graser	Austrian Institute of Technology (Austria)
Mahmoud Sakr	Université libre de Bruxelles (Belgium)
Yannis Theodoridis	University of Piraeus (Greece)

Program Committee:

Gennady Andrienko	IAIS Fraunhofer (Germany)
Alexander Artikis	University of Piraeus and NCSR Demokritos (Greece)
Somayeh Dodge	University of California Santa Barbara (USA)
Christos Doulkeridis	University of Piraeus (Greece)
Cong Gao	Nanyang Technological University (Singapore)
Gyözö Gidofalvi	KTH (Sweden)
Ioannis Kontopoulos	Harokopio University and NCSR Demokritos (Greece)
Hua Lu	Aalborg University (Denmark)
Mirco Nanni	ISTI-CNR (Italy)
Kjetil Nørvåg	Norwegian University of Science and Technology (Norway)
Kostas Patroumpas	Athena RC (Greece)
Nikos Pelekis	University of Piraeus (Greece)
Alessandra Raffaetà	Universita' Ca' Foscari Venezia (Italy)
Chiara Renso	ISTI-CNR (Italy)
Giulia Rovinelli	Universita' Ca' Foscari Venezia (Italy)
Marta Simeoni	Universita' Ca' Foscari Venezia (Italy)
Amilcar Soares	Linnaeus University (Sweden)
Panagiotis Tampakis	University of Southern Denmark (Denmark)
Konstantinos Tserpes	Harokopio University of Athens (Greece)
Karine Zeitouni	University of Versailles Saint-Quentin (France)

Data Systems Education (DataEd)

Data systems education is foundational in a variety of programs such as computer science, data science, and information systems and science. And, indeed, data management concepts are both timely and timeless in our increasingly data-driven world. A continual focus since the 1970's in the database research community is the place in curricula and best practices for teaching data systems concepts. This important conversation is particularly lively in recent years given the rise of data science. There is also a long tradition in the Computer Science Education research community on investigations into how students learn data systems concepts. With the increasing focus on data in the past decade, there is renewed focus on data systems in education research.

Both the DB and CS Education communities, and adjacent communities, e.g., in Statistics Education, have complementary perspectives and experiences to share with each other. There is much to be gained by bringing the communities more closely together: to share findings, to cross-pollinate perspectives and methods, and to shed light on opportunities for mutual progress in data systems education. Under the DataEd Initiative umbrella, we aim to organize various events for these communities to come together, for presentation and discussion of data management systems education experiences and research.

This year, the DataEd workshop is in its fifth iteration. We opened the floor to seven interesting papers, and two keynotes that bridged the research areas of CS Education and Data Systems research. We'd like to thank our keynote speakers prof. Craig Zilles from the University of Illinois at Urbana-Champaign, and prof. Azza Abouzied from New York University at Abu Dhabi.

Program Committee Chairs:

- Abdussalam Alawini — University of Illinois Urbana-Champaign (USA)
- George Fletcher — Eindhoven University of Technology (Netherlands)
- Daphne Miedema — University of Amsterdam (Netherlands)

Program Committee:

- Alan Fekete — University of Sydney (Australia)
- Giovanna Guerrini — University of Genova (Italy)
- Martin Goodfellow — University of Strathclyde (Scotland)
- Paul Groth — University of Amsterdam (Netherlands)
- Oscar Karnalim — Maranatha Christian University (Indonesia)
- Hui Li — Xidian University (China)
- Andrew Petersen — University of Toronto (Canada)
- Rachel Pottinger — The University of British Columbia (Canada)
- Stefanie Scherzinger — Universität Passau (Germany)
- Toni Taipalus — Tampere University (Finland)
- Thomas Zeume — Ruhr-Universität Bochum (Germany)

Transforming Graph Data (TGD)

Graphs are widely used to model interconnected real-world entities, requiring efficient storage, processing, and analysis. While a diverse ecosystem of graph database systems has emerged to tackle these challenges, graph transformation mechanisms remain underdeveloped. Critical gaps include the lack of formal frameworks for defining and applying graph transformations, as well as the absence of expressive syntactic and semantic primitives for querying temporal properties such as timeliness and versioning. Additionally, advancing interoperability, reliability, scalability, and adaptive learning in graph transformation ecosystems demands new models, techniques, and a deeper exploration of generative AI's role in automating and optimising these processes.

The TGD workshop received nine high-quality submissions, of which seven were accepted, including three shepherded papers, leading to an acceptance rate of 77%. The selected contributions underwent a rigorous review process and represent a diverse and timely cross-section of current research on graph transformations and their applications.

The accepted papers advance graph transformation and analysis along multiple dimensions. They include an optimization-based, quantum-inspired method for property graph schema discovery, and heterogeneity-aware graph data profiling to support cost-aware schema evolution and transformation. Additional contributions address explainable Datalog-based transformations with aggregation and LLM-assisted generation of explicit schema mappings for scalable RDF construction. Hybrid neuro-symbolic approaches are explored for query-driven knowledge graph summarization and inductive link prediction for data lineage discovery, while temporal-aware adversarial techniques shed light on robustness challenges in dynamic graph learning. Collectively, these works underscore the depth and evolving scope of research on scalable, explainable, and intelligent graph transformations.

In addition, the workshop featured an academic keynote and an industrial one.

Angela Bonifati (University Lyon 1, CNRS LIRIS, IUF, France) presented *"Property Graph Transformations in Action: From Data Integration to Causal Analysis"*. The keynote highlighted scalable declarative approaches to property graph transformations, with a focus on data integration and data cleaning, and explained how these techniques extend to causal inference and path-based causal analysis. The talk emphasized how property graphs can serve as a powerful integration paradigm and connect transformation techniques with emerging standards such as GQL and SQL/PGQ, as well as future schema and constraint languages.

Efthymia Tsamoura (Huawei Labs, Cambridge, UK) presented *"Trigger Graphs and Probabilistic Equivalence: Towards Scalable and Efficient Neurosymbolic Learning and Inference"*. The talk introduced trigger graphs, a scalable symbolic reasoning technique enabling exact Datalog reasoning over billion-edge graph stores within seconds. She also presented the new equivalence semantics for probabilistic logic programs, which improved neurosymbolic learning and inference, such as link prediction and rule mining, by up to 42% compared to state-of-the-art approaches.

The program concluded with a discussion on *"The Future of Graph Transformations",* bringing together keynote speakers, organizers, and participants to explore open challenges and next steps.

The TGD organisers extend their heartfelt thanks to everyone who contributed to its success:

- The authors, for their valuable research contributions and enriching discussions.
- The keynote speakers, Angela Bonifati and Efthymia Tsamoura, for honouring us with their presence and for inspiring new directions in graph data transformation research.
- The Program Committee members, for their diligent efforts in providing constructive and valuable feedback to authors.
- The EDBT/ICDT 2026 workshop and general chairs for their trust, patience, and guidance.

The second edition of the TGD workshop would not have been possible without the support of everyone involved. We are deeply grateful for this success and look forward to its continuation.

Program Committee Chairs:

- Anna Bernasconi — Politecnico di Milano (Italy)
- Stefania Dumbrava — ENSIIE, INRIA, IRIF, Télécom SudParis (France)
- Riccardo Tommasini — INSA Lyon (France)

Program Committee:

- Nelly Barret — INSA Lyon & CNRS LIRIS (France)
- Francesca Bugiotti — CentraleSupélec & CNRS LISN (France)
- Rachid Echahed — CNRS LIG, Grenoble Alpes University (France)
- Mauro Famà — INSA Lyon (France)
- Meike Klettke — Regensburg University (Germany)
- Haridimos Kondylakis — FORTH-ICS, University of Crete (Greece)
- Sergio Lifschitz — PUC-Rio (Brazil)
- Matteo Lissandrini — University of Verona (Italy)
- Victor Marsault — CNRS LIGM, Gustave Eiffel University (France)
- Andrea Mauri — Lyon 1 University, CNRS LIRIS (France)
- Marco Mesiti — University of Milano (Italy)
- Jyrki Nummenmaa — Tampere University (Finland)
- Alexandra Rogova — Warsaw University (Poland)
- Yuya Sasaki — Osaka University (Japan)
- Petra Selmer — Bloomberg (United Kingdom)
- Christopher Spinath — Lyon 1 University, CNRS LIRIS (France)
- Uta Störl — University of Hagen (Germany)

Explainable Data Science and Machine Learning for the Sciences (XAI4Science)

Over the last couple of decades, the increasing availability of advanced computational resources and big scientific data boosted data-driven methods in scientific discovery and innovation. From neuroscience and astrophysics, to medicine and pharmaceutics, chemistry and material sciences up to weather and climate sciences, scientists currently process large volumes of experimental data and employ data science and machine learning techniques to validate and generate scientific hypotheses. Unfortunately, existing AI systems used to engineer and analyse data are mainly opaque, i.e., it is difficult to understand why they return a specific output or what they could return if input data were slightly different. They typically make automated decisions by fixating on a particular hypothesis under investigation without providing evidence for or against it.

Recent advances in explainable artificial intelligence (XAI) aim to bridge the gap between human cognitive decision-making processes and AI systems. However, XAI methods mainly focus on understanding AI model behavior rather than on how to exploit it for discovering new human knowledge. Their impact in complex problem solving is currently limited by the lack of completeness, robustness, and universality across AI models, data modalities, and scientific pipelines.

In this scenario, we are glad to organize the XAI4Science workshop, which on the 24th of March, in conjunction with the EDTB conference, aims to bring together researchers, practitioners, and domain experts working at the intersection of data science, machine learning, and scientific disciplines to discuss advances in XAI methods that can effectively and efficiently support scientific discovery.

We have received eleven (11) high-quality submissions and accepted six (6) for presentation at the venue (three (3) long papers and three (3) short papers).
The workshop kick-off has been followed by a keynote presentation from Professor Giovanni Stilo (Luiss Guido Carli University, Luiss Business School) on Advances and Future Perspectives in Graph Counterfactual Explanations. The talk offered an overview of the field by introducing the conceptual foundations of GCE, describing the main families of explainers, and reviewing recent progress that spans perturbation-based approaches, global reasoning methods, dynamic-graph counterfactuals, and latent or spectral generative models. Then, Professor Giovanni Stilo's talk provided practical tools and a visual comparison of representative techniques. The session concluded with a forward-looking discussion that highlighted emerging research paths and open questions likely to shape the next phase of counterfactual explainability for graph-based learning.

Program Committee Chairs:

- Vassilis Christophides — ETIS, ENSEA (France), CNRS IPAL (Singapore)
- Jin-Song Dong — National University of Singapore (Singapore)
- Nicolas Labroche — Univ. of Tours, LIFAT (France)
- Evaggelia Pitoura — Univ. of Ioannina, Archimedes Athena RC (Greece)
- Céline Robardet — INSA Lyon, LIRIS (France)
- Yongfeng Zhang — Rutgers University (USA)

Program Committee:

- Julien Aligon — Univ. of Toulouse Capitole, IRIT (France)
- Alexandre Chanson — Univ. of Tours, LIFAT (France)
- Emmanuel Doumard — Univ. of Tours, LIFAT (France)
- Leilani Gilpin — Univ. of California Santa Cruz, AIEA (USA)
- Riccardo Guidotti — Univ. of Pisa, KDD (Italy)
- Moncef Garouani — Univ. Toulouse Capitole, IRIT Lab (France)
- Matthijs van Leeuwen — Leiden University, LIACS (Netherlands)
- Michele Linardi — ETIS, CNRS, CYU (France)

- Marie-Jeanne Lesot — Sorbonne Université / LIP6 (France)
- Patrick Marcel — Univ. of Orléans, LIFO (France)
- Christophe Marsala — Sorbonne Univ., LIP6 (France)
- Eirini Ntoutsi — Univ. of Bundeswehr Munich, CODE (Germany)
- Guillaume Renton — ETIS, CNRS, ENSEA (France)
- Konstantinos Stefanidis — Tampere Univ. (Finland)
- Giovanni Stilo — Luiss Business School (Italy)
- Simone Stumpf — Univ. of Glasgow (UK)
- Aikaterini Tzompanaki — ETIS, CNRS, CYU (France)
- Juntao Tan — Rutgers University (USA)

Quality in Large Language Models and Knowledge Graphs (QuaLLM-KG)

The rapid progress of knowledge graphs (KGs) and large language models (LLMs) has transformed data science, data management, and artificial intelligence. Knowledge graphs have become a central paradigm for representing structured information, while LLMs provide unprecedented capabilities in natural language understanding and generation.

Their success depends heavily on the quality of the underlying data, models, and processes. KGs suffer from incompleteness, incorrect links, outdated facts, and lack of provenance; LLMs suffer from hallucinations, bias, lack of factual grounding, and reproducibility issues.

When KGs and LLMs are combined—KG-grounded LLMs or LLM-assisted KG construction—the quality challenges multiply and require novel evaluation, integration, and trustworthiness solutions.

The workshop brings together communities across databases, AI, NLP, knowledge representation, information retrieval, data quality, and data governance to foster research on ensuring quality in KGs, LLMs, and their hybrid systems.

QuaLLM-KG aims to bridge theoretical, methodological, and empirical perspectives while providing demonstrations and practical case studies to give attendees hands-on understanding of real-world solutions for LLM/KG quality.

In this first edition, we received 9 papers (6 long and 3 short), among which 5 have been accepted (3 long and 2 short) which corresponds to an acceptance rate of 55%.

Program Committee Chairs:

- Soror Sahri — Université Paris Cité (France)
- Sven Groppe — University of Lübeck (Germany) and TU Bergakademie Freiberg (Germany)
- Farah Benamara — University of Toulouse (France) & CNRS IPAL (Singapore)

Program Committee:

- Asma Abboura — University of Hassiba Benbouali (Algeria)
- Hanieh Khorashadizadeh — University of Lübeck (Germany)
- Hazar Harmouch — University of Amsterdam (Netherlands)
- Imane Hocine — University of Luxembourg (Luxembourg)
- Jesualdo Fernandez Breis — Universidad de Murcia, IMIB-Arrixaca (Spain)
- Mohammad Sadoghi — University of California (USA)
- Mourad Ouzzani — Qatar Computing Research Institute, HBQU (Qatar)
- Nathalie Aussenac — IRIT–CNRS (France)
- Raphaël Troncy — EURECOM (France)
- Rinaldo Lima — Federal Rural University of Pernambuco (Brazil)
- Sanju Tiwari — Sharda University (India)
- Valter Uotila — University of Helsinki (Finland)