

DataMod2020: 9th International Symposium "From Data to Models and Back"

Juliana Bowles
jkfb@st-andrews.ac.uk
University of St Andrews
St Andrews, UK

Giovanna Broccia
giovanna.broccia@isti.cnr.it
FMT Lab, ISTI-CNR
Pisa, Italy

Mirco Nanni
mirco.nanni@isti.cnr.it
KDD Lab, ISTI-CNR
Pisa, Italy

ABSTRACT

DataMod 2020 aims to bring together practitioners and researchers from academia, industry and research institutions interested in the combined application of computational modelling methods with data-driven techniques from the areas of knowledge management, data mining and machine learning. Modelling methodologies of interest include automata, agents, Petri nets, process algebras and rewriting systems. Application domains include social systems, ecology, biology, medicine, smart cities, governance, security, education, software engineering, and any other field that deals with complex systems and large amounts of data. Papers can present research results in any of the themes of interest for the symposium as well as application experiences, tools and promising preliminary ideas. Papers dealing with synergistic approaches that integrate modelling and knowledge management/discovery or that exploit knowledge management/discovery to develop/synthesise system models are especially welcome.

CCS CONCEPTS

• **Computing methodologies** → **Model development and analysis**; **Machine learning**; • **Information systems** → **Data mining**.

ACM Reference Format:

Juliana Bowles, Giovanna Broccia, and Mirco Nanni. 2020. DataMod2020: 9th International Symposium "From Data to Models and Back". In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM '20)*, October 19–23, 2020, Virtual Event, Ireland. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3340531.3414073>

1 ABOUT THE WORKSHOP SERIES

The DataMod Symposium series started with the first two editions as "International Symposium on Modelling and Knowledge Management for Sustainable Development" (MoKMaSD), both associated with SEFM in 2012 and 2013. In the 2014 edition, the scope of the symposium was expanded by removing the constraint on the considered application domain. Consequently, the title of the symposium was changed to "International Symposium on Modelling and Knowledge Management applications: Systems and Domains" (with the same acronym). This development of the MoKMaSD Symposia was the result of the discussions held during the first two editions. In

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CIKM '20, October 19–23, 2020, Virtual Event, Ireland

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6859-9/20/10.

<https://doi.org/10.1145/3340531.3414073>

particular, from such discussions it emerged that the synergistic application of modelling and knowledge management (and discovery) would have been an interesting topic in itself, independently of the application domain.

The 2014 edition of MoKMaSD was a sort of "exploratory" edition aimed at building a community around the topic of the integrated use of management/discovery methodologies and modelling. The result of such an "exploration" was that 7 full papers were accepted, 4 of which from scientists mainly working on modelling methodologies and 3 of which from scientists mainly working on knowledge management/discovery methodologies. Some of the papers contained seeds of synergistic application of such two kinds of methodologies. Moreover, during the Symposium, the potential synergies became the subject of discussions both during the presentation of the regular papers and during the presentations by the two invited speakers.

In 2015, MoKMaSD started receiving papers proposing new synergistic modelling/data mining approaches. Hence, the symposium started giving the results we expected. Moreover, the success of the symposium continued to increase, with 13 submitted papers, 9 of which were accepted for publication in an LNCS post-proceedings.

In 2016, with the new name DataMod, the number of submissions continued to increase. Moreover, for the first time after the symposium, a special issue of the Journal of Intelligent Information Systems (JIIS) was organised, in order to collect extended and revised versions of DataMod papers as well as new submissions on topics of the symposium. Furthermore, we created a Steering Committee for the symposium that included some of the organisers of the previous editions (who had contributed to the new shape of the symposium) as well as international experts on data mining, knowledge management and computational modelling. We believe that the Steering Committee gave the symposium more solid foundations and a guarantee of stability for the following years.

In 2017, DataMod continued to grow, becoming a two-day event. As in 2016, a journal special issue (this time of the Journal of Algebraic and Logical Methods in Programming) was organised.

In 2018, the 7th edition of DataMod was associated with the International Conference on Software Engineering and Formal Methods (SEFM), and the 8th edition of DataMod in 2019 with the 3rd World Congress on Formal Methods (FM).

Over the years DataMod has mainly been collocated with conferences such as SEFM and FM, which attract a community with an interest in modelling and formal methods. Despite the evident increasing success that DataMod has had over the years bringing together researchers working on various data-driven application domains, there is scope to expand the interests and contributions

made to DataMod further. Indeed, by running it as a satellite event of CIKM 2020 we are hoping to establish new research connections with a further community with a stronger focus on information retrieval and knowledge management. We believe that this will increase the success and widen the interest of DataMod in the future and promote discussions across communities.

2 THEME AND TOPICS

DataMod has a wide range of themes and topics. Considered modelling and analysis methodologies include: *Agent-based Methodologies, Automata-based Methodologies, Big Data Analytics, Cellular Automata, Classification, Clustering, Segmentation and Profiling, Conformance Analysis, Constraint Programming, Data Mining, Differential Equations, Game Theory, Machine Learning, Membrane Systems, Network Theory and Analysis, Ontologies, Optimisation Modelling, Petri Nets, Process Calculi, Process Mining, Rewriting Systems, Spatio-temporal Data Analysis/Mining, Statistical Model Checking, Text Mining, and Topological Data Analysis.*

There is a particular interest in submissions from various application domains such as: *Biology, Brain Data and Simulation, Business Process Management, Climate Change, Cybersecurity, Ecology, Education, Environmental Risk Assessment and Management, Enterprise Architectures, Epidemiology, Genetics and Genomics, Governance, HCI and Human Behaviour, Open Source Software Development and Communities, Pharmacology, Resilience Engineering, Safety and Security Risk Assessment, Social Good, Social Software Engineering, Social Systems, Sustainable Development, Threat modelling and analysis, Urban Ecology, Smart Cities and Smart Lands.*

Synergistic approaches may include:

- (1) the use of modelling methods and notations in a knowledge management/discovery context,
- (2) the development and use of common modelling and knowledge management/discovery frameworks to explore and understand complex systems from the application domains of interest.

3 WORKSHOP ORGANISERS

All contributions in the form of either regular papers (up to 15 pages) or short papers (up to 8 pages) is reviewed by three Program Committee members, and is evaluated on the basis of originality, contribution to the field, technical and presentation quality, and relevance to the symposium. The members of the 9th International Symposium “From Data to Models and Back” are:

- Oana Andrei, University of Glasgow, UK
- Davide Basile, ISTI-CNR, Italy
- Mario Boley, Monash University, Australia
- Juliana Bowles, University of St Andrews, UK (Co-Chair)
- Giovanna Broccia, ISTI-CNR, Italy (Co-Chair)
- Marco Caminati, University of St Andrews, UK
- Antonio Cerone, Nazarbayev University, Kazakhstan
- Ricardo Czekster, Newcastle University, UK
- Flavio Ferrarotti, SCCH, Austria
- Lars Kotthoff, University of Wyoming, USA
- Giulio Masetti, ISTI-CNR, Italy
- Sotiris Moschoyannis, University of Surrey, UK
- Paolo Milazzo, University of Pisa, Italy

- Anna Monreale, University of Pisa, Italy
- Reshma Munbodh, Brown University, USA
- Mirco Nanni, ISTI-CNR, Italy (Co-Chair)
- Lucia Nasti, University of Pisa, Italy
- Céline Robardet, INSA Lyon, France

The program co-chairs are: Juliana Bowles (Computer Science, University of St Andrews), Giovanna Broccia (FMT Lab, ISTI-CNR, Italy), and Mirco Nanni (KDD Lab, ISTI-CNR, Italy).

The organizing committee is composed by: Oana Andrei (University of Glasgow, UK), Antonio Cerone (Nazarbayev University, Kazakhstan), and Paolo Milazzo (University of Pisa, Italy).

The steering committee is composed by:

- Oana Andrei, University of Glasgow, UK
- Antonio Cerone, Nazarbayev University, Kazakhstan
- Vashti Galpin, University of Edinburgh, UK
- Riccardo Guidotti, University of Pisa, Italy
- Marijn Janssen, Delft University of Technology, the Netherlands
- Stan Matwin, University of Ottawa, Canada
- Paolo Milazzo, University of Pisa, Italy
- Anna Monreale, University of Pisa, Italy
- Mirco Nanni, ISTI-CNR, Italy

4 ACCEPTED PAPERS

Accepted papers of DataMod 2020 will be published as a dedicated volume of Lecture Notes in Computer Science, Springer, as a post-proceedings after the workshop.

DataMod 2020 received a mix of papers from different application domains, with a strong focus on healthcare this year. The papers combine a variety of approaches from process mining, text mining, machine learning, deep learning, discrete event simulation, constraint programming, formal modelling and big data, in order to improve the analysis of epidemic models, demand prediction, scheduling, network optimisations, gain an understanding of network behaviour, improve data integration, emulate knowledge generation, gain knowledge from medical text, and so on.

Submitted papers came from 12 different countries: Brazil, Bulgaria, China, France, Germany, India, Italy, Kazakhstan, Tunisia, UAE, UK and USA.

5 INVITED SPEAKER

The invited speaker at DataMod 2020 is Michael Vinov, head of the Security and Data Quality group at the IBM Haifa Research Lab.