Data & Knowledge Engineering xxx (xxxx) xxx

Contents lists available at ScienceDirect

Data & Knowledge Engineering

journal homepage: www.elsevier.com/locate/datak

Preface

Bridging conceptual modeling and machine learning

This special issue of *Data & Knowledge Engineering* brings together extended and revised versions of a selection of the best papers presented at the International Conference on Conceptual Modeling (ER) 2024, held at Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, USA, during October 28–31, 2024. The papers included here were nominated by the conference's program committee on the basis of their originality and technical quality. They were subsequently invited for submission to this special issue, where they underwent substantial extension and an independent round of peer review before final acceptance. Together, they offer a timely snapshot of a field that is actively redefining its boundaries in response to the growing prominence of machine learning and artificial intelligence.

Conceptual modeling has, for several decades, provided the discipline of information systems with a principled way to represent the semantics of a domain, including its entities, relationships, processes, and goals, independent of any particular implementation. Machine learning, by contrast, has largely developed along a different trajectory, one driven by data and statistical inference rather than explicit domain semantics. As machine learning systems have moved from research prototypes into everyday operational use. However, the limitations of purely data-driven approaches have become increasingly apparent: models that are difficult to interpret, pipelines that are hard to govern, and data preparation processes that lack principled documentation. Conceptual modeling offers a complementary set of tools and techniques precisely where these limitations are most keenly felt. This special issue is motivated by the conviction that the two fields have a great deal to offer one another.

The relationship between conceptual modeling and machine learning runs in both directions, and the contributions in this issue reflect that duality. In one direction, conceptual modeling *supports* machine learning: ontologies and domain models can guide data preparation and feature engineering, constrain and document model development, and provide the semantic scaffolding needed for explainable AI. In the other direction, machine learning systems themselves become *objects of* conceptual modeling: their pipelines, lifecycles, and engineering artifacts can be represented, analyzed, and reasoned about using the same modeling techniques traditionally applied to enterprise systems and business processes. A third thread that runs through this bridge concerns responsibility: the ethical, societal, and legal implications of AI systems, and the extent to which conceptual models can help operationalize requirements such as fairness, accountability, and transparency at the design stage rather than after deployment. This expansion mirrors a broader shift within the conceptual modeling community's own flagship venues, where calls for papers now explicitly invite work at the intersection of modeling and AI alongside the field's traditional concerns.

The papers selected for this special issue were not chosen in isolation from that traditional core. Rather, they extend established lines of conceptual modeling research, including work on ontologies, process modeling, and goal modeling, into contact with machine learning, rather than treating the two as separate concerns. Authors of nominated papers were invited to prepare substantially extended versions of their conference contributions, incorporating new results, expanded evaluations, and deeper engagement with related work, consistent with the journal's requirements for conference-linked special issues. Each extended submission then underwent a full round of independent peer review, distinct from the review process at the originating conference, with final decisions made by the guest editors in consultation with the reviewers.

This special issue collects three papers, each illustrating a different facet of the relationship between conceptual modeling and machine learning, and more broadly between conceptual modeling and the engineering of intelligent, interactive systems.

Two of the papers extend goal-oriented modeling, a well-established conceptual modeling paradigm, into new territory. Sotirios Liaskos, John Mylopoulos, Alex Borgida, and Shakil M. Khan, in "Goal-Oriented Modeling and Analysis of Explanation Requirements," address a core concern of explainable AI: before a system can explain its actions, the requirements for that explanation must themselves be systematically captured and analyzed. The authors propose a framework in which explanation requirements are represented as

https://doi.org/10.1016/j.datak.2026.102637

Available online 5 August 2026

stakeholder goals and refined into explanation tasks and interaction templates, extended with temporal and causal constraints so that the resulting model can generate answers to stakeholder questions. An empirical analysis based on two system cases shows that the explanations generated in this way were rated as useful by the participants who took on the role of recipients of the explanations.

In "Designing and Generating Reinforcement Learning Training Simulators using Goal Models," Sotirios Liaskos, Nina Dang, Shakil M. Khan, John Mylopoulos, Reza Golipour, and Amirreza Radjou turn goal modeling toward a different challenge: the construction of simulation environments for training reinforcement learning agents. Rather than treating simulator development as an ad hoc implementation task, the authors introduce extensions to a standard goal modeling notation that allow an agent's possible interactions with its environment to be modeled concisely, then formalized and executed by a purpose-built simulation engine that off-the-shelf RL algorithms can train against. Along with the first paper, this contribution shows how a shared conceptual modeling paradigm, goal orientation, can support machine learning systems from two very different angles: explaining their behavior after the fact, and specifying the environments in which their behavior is learned in the first place.

The third paper broadens the issue's scope from machine learning specifically to the wider landscape of conceptual modeling for emerging technology domains. In "ARWFML: A Domain-Specific Modeling Language for Augmented Reality Applications," Fabian Muff and Hans-Georg Fill address the complexity of building augmented reality (AR) software, which traditionally demands advanced coding skills, by proposing a domain-specific visual modeling language for designing AR scenarios as executable applications. Developed through four design science research cycles, including the creation of a new web-based 2D and 3D modeling environment (MM-AR), the language allows AR workflows and augmentations to be defined graphically and executed directly, demonstrating that conceptual modeling's traditional strength of raising the level of abstraction for complex system development extends naturally to immersive and interactive technologies as well.

Together, these three papers show a field extending its established methods, most notably goal-oriented modeling, into domains where explicit semantics are urgently needed: explaining the actions of autonomous systems, engineering the training environments of learning agents, and lowering the barrier to building increasingly complex interactive applications. Two threads run through the collection. First, goal-oriented modeling appears to be an unusually versatile bridge between conceptual modeling and machine learning, valuable in both a retrospective role, explaining behavior, and a prospective one, specifying training environments. Second, the discipline's core methodological strength, providing higher-level, tool-supported abstractions over complex technical systems, remains as relevant to today's AI and immersive technologies as it has been to the enterprise and information systems that conceptual modeling has traditionally served. Considerable work remains to connect these threads further, for instance by investigating whether models used to specify a system's training environment might also serve as a basis for later explaining its behavior. Nevertheless, the papers in this issue provide a solid foundation for that future work.

We are grateful to the authors for their contributions and for their willingness to extend their work for this special issue, to the reviewers who generously gave their time and expertise to the evaluation of the submissions, to the program chairs of ER 2024 for their support in identifying and nominating candidate papers, and to the Editor-in-Chief and editorial office of *Data & Knowledge Engineering* for their guidance throughout the process.

The papers in this issue represent an early, but substantive, step in bringing conceptual modeling and machine learning into closer dialogue. We hope they will encourage further work at this intersection, and that the conceptual modeling community will continue to find in machine learning both a challenge worth addressing and a set of tools worth adopting. In turn, we hope the machine learning community will come to see conceptual modeling as a natural partner in building relevant and right AI systems that are more transparent, more governable, and more trustworthy.

Wolfgang Maass[*], Hyoil Han, Hasan Yasar

[*] Corresponding author.
E-mail address: wolfgang.maass@iss.uni-saarland.de (W. Maass).