Opportunities and Challenges for GORE in the era of Learning-based AI

iStar 2024 Panel Report

Travis D. Breaux[1], Giancarlo Guizzardi[2], Eric Yu[3], Amal Ahmed Anda[4], Sotirios Liaskos[5] and Elda Paja[6]

[1]*Carnegie Mellon University, USA*
[2]*University of Twente, The Netherlands*
[3]*University of Toronto, Canada*
[4]*University of Ottawa, Canada*
[5]*York University, Canada*
[6]*IT University of Copenhagen, Denmark*

Abstract

The following discussion paper summarizes the results of a panel discussion conducted on October 28, 2024 at the 17th iStar International Workshop, co-located with the International Conference in Conceptual Modelling (ER) in Pittsburgh, PA, United States. The panelists included Travis D. Breaux from Carnegie Mellon University, US, Giancarlo Guizzardi from the University of Twente, The Netherlands, and Eric Yu from the University of Toronto, Canada. The panel was moderated by Elda Paja from the IT University of Copenhagen, Denmark.

Keywords

iStar, goal-oriented modelling, learning-based AI, ML-based applications

1. Introduction

The emergence of learning-based AI systems has presented significant challenges and opportunities in various domains, from content generation to decision making. The iStar 2024 panel "Opportunities and Challenges for Goal-oriented Requirements Engineering (GORE) in the era of Learning-based AI" focused on the respective challenges and opportunities of such systems in the area of Goal Oriented Requirements Engineering (GORE), and conceptual modeling in general. The panelists, Travis Breaux, of Carnegie Mellon, Giancarlo Guizzardi of the University of Twente, and Eric Yu from the University of Toronto, were asked about the role of conceptual modeling in general and goal modeling specifically in safeguarding the quality and enhancing the performance of learning-based AI systems, the impact of such systems in the general software development and requirements engineering process, and the motivation for studying "traditional" knowledge representation-based methods, including conceptual modeling, when deep-learning seems to monopolize attention both of the society at large and of students and future engineers.

The following sections provide a summary of the panel discussion, highlighting key disagreements, as well as the points where the panelists found common ground.

iStar'24: The 17th International i Workshop, October 28, 2024, Pittsburgh, PA, USA
breaux@cs.cmu.edu (T. D. Breaux); g.guizzardi@utwente.nl (G. Guizzardi); eric.yu@utoronto.ca (E. Yu); amal_eletri@yahoo.com (A. A. Anda); liaskos@yorku.ca (S. Liaskos); elpa@itu.dk (E. Paja)
https://www.cs.cmu.edu/~breaux/ (T. D. Breaux); https://www.giancarloguizzardi.com/ (G. Guizzardi); http://www.cs.toronto.edu/~eric/ (E. Yu); https://www.yorku.ca/liaskos/ (S. Liaskos); http://itu.dk/~elpa/ (E. Paja)
0000-0001-7127-8155 (T. D. Breaux); 0000-0002-3452-553X (G. Guizzardi); 0000-0002-3832-5981 (E. Yu); 0000-0001-7851-0199 (A. A. Anda); 0000-0001-5625-5297 (S. Liaskos); 0000-0002-8346-2467 (E. Paja)

2. On the Necessity of Goals

2.1. Goals, values, and technologies

The panel first discussed the opportunities and challenges for goal modeling in developing and evaluating learning-based AI systems. The notion of value came up, as, roughly, a degree to which certain capacities of an artifact contribute to the satisfaction of an agent's desires [1]. This value dimension can be argued to be orthogonal to the technical characteristics of the artifact. Whatever the latter is (e.g., traditional vs. learning-based systems) investigation of its value vis-a-vis stakeholder needs is always pertinent.

The fact that learning-based systems are data-driven does not diminish the relevance and role of goals and values. On the contrary, the panel claimed, critical components of AI system design such as selecting features, judging data quality, must align with the system's broader goals to ensure it operates fairly and effectively. As demonstrated, e.g., in [2], the notion of value (and risk as its dual opposite notion) are fundamental for understanding ethical dimensions such as beneficence, non-maleficence, and explicability. Moreover, as briefly discussed in [3], it could also be used to analyze the notion of fairness. In other words, without a clear theory of value, it is impossible to assess whether an AI system functions correctly, behaves ethically or is biased in some way.

2.2. The blind audition case

The metaphor of blind auditions was introduced to illustrate the relevance of a theory of value for fairness. For example, take: fairness to be thought as 'treating in an equivalent way entities that are equivalent under value assessment' [3]; value assessment to select which aspects (capacities, qualities) of the value object contribute to a given set of goals [1, 2]. By explicitly asserting goals that should be considered in a given context, we can derive which features are important (have more value) in that context (i.e., music rendition ability), against features that are unimportant, add noise, or whose consideration could even be considered harmful, (e.g., gender, ethnicity, age). The point is that we are always confronted with such decisions when purposefully developing artifacts, and goals are the source of the criteria we use to make the decisions. That the artifacts themselves are based on a new kind of technology is orthogonal to value and purpose.

3. The challenges in explicating and evaluating goals

Despite the strong case for the relevance of goals when analyzing learning-based AI systems, the panel identified challenges in extracting and articulating such goals, as well as evaluating a system with respect to whether it meets such goals. Traditional metrics like accuracy, precision, and recall are commonly used in machine learning to evaluate task-specific output. However, these metrics are less effective when the output is generative and may not directly align with well-defined, quantifiable goals. In many contexts, such as writing emails or generating travel plans, users may not explicitly define their goals. For example, when generating an email, the user may not clearly state their objective, making it difficult to assess whether the generated text meets the goal. Similarly, in generative tasks such as travel planning, the AI may produce a plan that lacks essential factors (e.g., travel time between destinations), which only becomes apparent to the user when she encounters issues in real-world scenarios. The inability to take into account these tacit requirements is related to the lack of capacity of current AI systems to generally reason with common-sense theories [4]. This makes the case for more intensive effort to discover and explicate goals so that they can be promptly used for the development/training of the system. The problem seems to also be domain specific, were goal structures and satisfaction in, e.g., the travel planning domain are not usable in the medical domain.

4. Humans, Machines, and Socio-technical Systems

4.1. Learning-based systems and the foundational principles of RE

The panelists went on to discuss RE in the AI era by first recalling foundational principles of RE - that RE is fundamentally about relationships between the machine and its environment, and thus requires proper understanding and analysis of the problem domain [5, 6].

With the prevalence of data-driven and learning-based techniques, modern systems are becoming tightly engaged with the environment. In the past, the feedback path from empirical evaluation back to requirements and design was slow and tenuous. Today, this feedback for learning and redesign can be instant and constant. On one hand, machines are becoming more human-like. On the other hand, human behaviour is increasingly influenced and shaped by technology systems. This tight coupling suggests that requirements modeling today needs to be able to support analysis of the kinds of effects that machines and humans are having on each other in this age of AI.

4.2. The relevance of agent-orientation and socio-technical analysis

Once we recognize that machines today operate within more complex human social contexts than in the past, it is immediately apparent that the kind of agent-oriented social actor dependency analysis offered by i^* can be very relevant and useful. Consider how a software vendor might utilize AI in providing customer support for their products [7]. Customers may talk directly with an AI agent knowledgeable about the product, or they may talk with a human agent that consults AI. Or the support AI can be embedded in the product itself. Effectiveness of the support function would depend on how quickly and accurately the support person or AI could diagnose and resolve the problem while keeping the customer cool. How the support agent will be motivated, trained and evaluated are also important considerations. More generally, AI applications can have far reaching effects on environmental sustainability, privacy, security, reputation and trust, and even professional and personal identity.

4.3. The challenge of modeling the human social context

The iStar framework was originally inspired by agent-oriented concepts from the earlier AI paradigm for artificial agents, and had limited expressiveness for modeling the human social context [8, 9]. Nevertheless, by offering an actor abstraction that can equally be used to model humans and machines, i^* modeling avoids the pitfall of prejudging the human-machine boundary, an important principle for requirements analysis. This approach allows the task of distributing responsibilities among actors to fall to RE, based on the capabilities and qualities of the relevant actors, such as specific classes of human and automated actors [10, 11].

As the kinds of relationships between humans and machines have become much more complex in the age of AI, the challenge is how to enrich requirements modeling, such as i^* modeling, to encompass such characteristics as human values and emotions. Fairness, toxicity, ethics and deception are among societal issues and concerns that recent foundations models have raised. The panelists deliberated on how the bodies of knowledge in the human and social disciplines can help enrich requirements analysis. One approach would be to extract abstractions from those disciplines. A major hurdle is the difficulty of accommodating the great diversity of perspectives and theories among those disciplines. Another approach would be to start from well established conceptual modeling constructs. For example, one might consider whether current formulations of the part-whole relationship (e.g., [12]) are adequate when the parts and wholes are social actors with agency, intentionality, and autonomy, within a rich social context of trust, emotions, history, and culture. AI applications today seem to bring up the very notion of agent and agency as the subject of investigation for conceptual modeling.

Legal implications and accountability in the context of agency (law of agency vs. law of personhood) are also pertinent. It was noted that the law of personhood does not apply to machines. But there is still a question of how legal accountability will be attributed.

The panel discussed the potential perils of not clearly distinguishing between humans and artificial agents in modeling. Artificial agents may have the appearance but may lack the substance of the reasoning capabilities that humans ascribe to them. In addition we must be careful to not attribute agency to phenomena when no such agency exists; the pareidolia [13] metaphor was specifically mentioned here, i.e., we seem prone, when dealing with some of these systems, to attribute unwarranted levels of intentionality and reasoning capacity. It is, hence, preferable to be cautious when viewing AI technologies as agents in par with humans. It is perhaps preferable to see AI systems as tools, as an enabling mechanism.

There is also the question of when autonomy can be granted to artificial intelligence, something the panel was skeptical about. One viewpoint was that it is the task of requirements engineering to determine the degrees of freedom to be granted to various agents (or actors in i^*), human or otherwise.

4.4. Opportunities for the Modeling and RE processes

The panel was asked what opportunities there are for learning-based AI in the practice of RE. Panelists pointed out that indeed requirements in the industrial context is very expensive and automation can support areas such as innovation as well as acquiring and implementing highly personalized requirements.

A combination of the traditional top-down with the newer bottom-up approach to modeling was also put forth. In the earlier days of conceptual modeling top-down modeling based on analytical effort by modelers proved challenging due to the complexity of the world with all its exceptions and subtleties. The bottom-up approach that was adopted in response, however, soon led to the realization that concepts were needed for organizing and reasoning with the observed phenomena. Process models [14] and knowledge graphs [15] were mentioned as areas that have seen this shift of focus from purely top-down approaches to one in which combinations of top-down and bottom-up approaches are employed. So a possible avenue would be to build evidence-informed models that allow different kinds of contradictions and inconsistencies while using stable aprioristic principles to guide the process. In goal models, this would imply detecting goals form data – e.g., hypothesized goals or contradictory goals. But there is also the opportunity of agent orientation: asking who has the goals.

It was, however, pointed out that learning-based systems alone, including LLMs, may not be reliable reasoning agents to realize the bottom-up aspect. Rather, we may need various kinds of hybridization via, e.g., neuro-symbolic methods. In general, our progress in devising abstractions to tackle this problem seems to be lagging behind the technological developments. As an example of the challenge, DAML (the DARPA Agent Markup Language) was mentioned as an early effort to enable intelligent agents in the web; 25 years after its introduction, micro-service composition is still performed by humans. A reason may be that the agent concepts in 2000 were about machines and were lacking abstractions, such as e.g., the ones that i^* languages propose. AlphaFold [16] was also mentioned as a demonstration of the level of domain modeling and data curation work that needed to accompany the core learning process. There are other examples of famous AI successes that did not rely entirely on learning as is commonly believed, but are rather neuro-symbolic systems. This points both to the limitations of learning-based system, and to the role of modeling in enhancing and safeguarding those systems.

So where should effort be dedicated in addressing the AI challenge? Should we invest in building abstractions for requirements analysis that model both humans and machines, in addition to giving time to those who test the technologies in the field and see where their experimentation takes them? The panel was inconclusive, though agreed that there is certainly room for substantial work on this topic using a diversity of approaches.

5. Summary

The panel had an informative and thought-provoking discussion about the opportunities and challenges in applying learning-based AI technologies in the areas of conceptual modeling and requirements engineering. The necessity of value and, hence, goals as concepts that allow us to explore the purposes of these artifacts was acknowledged. Further, the vision of a generalized abstraction for modeling the full range of agents (from human to artificial) was put forth. It was countered with reservations about the actual capabilities and dependability of the AI systems in question, which has shown to be limited or unpredictable at best. This challenges the view that such systems can be currently modeled with the same concepts and tools as we would model humans. However ambitious, the project of a unifying modeling approach may nevertheless be a well-motivated investment for the future and the i^* family of languages may be a natural ground on which it can be based. At the same time, it may be useful to also be attentive to how these systems are used in the specific fields of application, and what is learned from those applications. The panel agreed that there is plenty of opportunity to, firstly, utilize these technologies for assisting the requirements engineering and modeling practice, including goal modeling – in, e.g., a combined top-down and bottom-up fashion – and, secondly, to use conceptual modeling in combination with learning-based systems to overcome the limitations and dangers of the latter. The latter opportunity is motivated by the recognition of the intrinsic limitations of pure learning-based systems as well as the realization that many famous AI success stories are based on neuro-symbolic approaches where modeling and learning from data are used synergistically.

References

[1] T. P. Sales, F. Baião, G. Guizzardi, J. P. A. Almeida, N. Guarino, J. Mylopoulos, The common ontology of value and risk, in: Conceptual Modeling: 37th International Conference, ER 2018, Xi'an, China, October 22–25, 2018, Proceedings 37, Springer, 2018, pp. 121–135.

[2] R. Guizzardi, G. Amaral, G. Guizzardi, J. Mylopoulos, An ontology-based approach to engineering ethicality requirements, Software and Systems Modeling 22 (2023) 1897–1923.

[3] D. Bork, I. David, S. España, G. Guizzardi, H. A. Proper, I. Reinhartz-Berger, The role of modeling in the analysis and design of sustainable systems: A panel report, Communications of the Association for Information Systems 54 (2024) 41.

[4] G. Marcus, E. Davis, Rebooting AI: Building artificial intelligence we can trust, Vintage, 2019.

[5] P. Zave, M. Jackson, Four dark corners of requirements engineerin, ACM Trans. Softw. Eng. Methodol. (TOSEM) 6 (1997) 1–30.

[6] A. Bennaceur, T. T. Tun, Y. Yu, B. Nuseibeh, Requirements engineering, in: Handbook of software engineering, Springer, 2019, pp. 51–92.

[7] E. Brynjolfsson, D. Li, L. R. Raymond, Generative AI at work (No. w31161), National Bureau of Economic Research, 2023.

[8] A. Newell, The knowledge level, Artificial intelligence 18 (1982) 87–127.

[9] M. Wooldridge, N. Jennings, Intelligent agents: Theory and practice, The knowledge engineering review 10 (1995) 115–152.

[10] A. Van Lamsweerde, Goal-oriented requirements engineering: A guided tour, in: Proceedings fifth ieee international symposium on requirements engineering, IEEE, 2001, pp. 249–262.

[11] E. Yu, P. Giorgini, N. Maiden, J. Mylopoulos, Social modeling for requirements engineering, MIT Press, 2010.

[12] G. Guizzardi, Ontological foundations for conceptual part-whole relations: the case of collectives and their parts, in: Advanced Information Systems Engineering: 23rd International Conference, CAiSE 2011, London, UK, June 20-24, 2011. Proceedings 23, Springer, 2011, pp. 138–153.

[13] M. Hamilton, S. Stent, V. DuTell, A. Harrington, J. Corbett, R. Rosenholtz, W. T. Freeman, Seeing faces in things: A model and dataset for pareidolia, in: European Conference on Computer Vision, Springer, 2024, pp. 377–395.

[14] W. Van Der Aalst, Process mining: Overview and opportunities, ACM Transactions on Management Information Systems (TMIS) 3 (2012) 1–17.

[15] E. Schultes, M. Roos, L. O. Bonino da Silva Santos, G. Guizzardi, J. Bouwman, T. Hankemeier,

A. Baak, B. Mons, Fair digital twins for data-intensive research, Frontiers in big data 5 (2022) 883341.

[16] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589. URL: https://doi.org/10.1038/s41586-021-03819-2. doi:10.1038/s41586-021-03819-2.