

Challenges Requiring the Combination of Machine Learning and Knowledge Engineering

Andreas Martin¹

¹FHNW University of Applied Sciences and Arts Northwestern Switzerland, School of Business, Riggenbachstrasse 16, 4600, Olten, Switzerland

Abstract

The AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering brought together researchers and practitioners from machine learning and knowledge engineering. The goal was to explore how combining these two fields can help address future AI challenges. The symposium included a joint keynote presentation by AI pioneers, over 25 presentations by contributors and authors who shared their research findings, and two challenges for the community to tackle in a follow-up event. This paper reports on the symposium and focuses on the current trend of generative AI and large language models (LLMs) and its possible synergy with knowledge-based systems (KBS), as the keynote speakers and the symposium chair emphasized. The discussions highlighted the potential of combining KBS's knowledge representation capabilities with LLMs' language generation capabilities.

1. Introduction

The AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering [1] brought together researchers and practitioners from machine learning and knowledge engineering domains. The symposium provided a platform to contemplate the potential contributions of merging these two fields in addressing a wide range of future challenges, encompassing societal, environmental, business, and fundamental AI challenges.

Notably, AAAI-MAKE 2023 marked the fifth edition of this symposium [2, 3, 4, 5], representing an evolving field that continues to adapt and respond to emerging trends and advancements in the intersection of machine learning and knowledge engineering. The AAAI-MAKE 2023 began with a remarkable highlight: a joint keynote presentation delivered by two distinguished AI pioneers and renowned experts, Edward A. Feigenbaum and Douglas B. Lenat.

Edward A. Feigenbaum, an esteemed professor emeritus of Computer Science at Stanford University, a recipient of the prestigious ACM Turing Award, and a Fellow of the AAAI, is often called the “father of expert systems.”

In A. Martin, K. Hinkelmann, H.-G. Fill, A. Gerber, D. Lenat, R. Stolle, F. van Harmelen (Eds.), Proceedings of the AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAAI-MAKE 2023), Hyatt Regency, San Francisco Airport, California, USA, March 27-29, 2023.


✉ andreas.martin@fhnw.ch (A. Martin)

🌐 <https://andreamartin.ch> (A. Martin)

🆔 0000-0002-7909-7663 (A. Martin)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

Meanwhile, Douglas B. Lenat, the visionary behind the Cyc project and founder of Cycorp, is a Fellow of the AAAI. During their captivating joint keynote presentation, they shared invaluable insights and perspectives on AI's historical milestones and future trajectory, shedding light on the challenges and opportunities of integrating machine learning and knowledge engineering. Additionally, they provided valuable commentary on the growing prevalence of generative AI and the trends surrounding large language models (LLMs), prompting thoughtful reflections within the symposium's audience.

2. Contributions and Challenges

The symposium featured over 25 presentations by contributors and authors, showcasing their papers, datasets, ontologies, and initial research findings. These presentations encompassed a wide array of topics, including but not limited to hybrid (human-artificial) intelligence and the concept of human-in-the-loop interactions. Moreover, the discussions delved into commonsense reasoning and explainable AI, highlighting the importance of imbuing AI systems with the ability to provide transparent and interpretable outputs. The symposium explored research directions, such as hybrid AI approaches and neuro-symbolic AI, which combine the strengths of both symbolic reasoning and neural networks. In addition, human-centered AI, dialogue systems, and conversational AI were explored, recognizing the significance of designing AI systems that effectively interact and communicate with humans. Lastly, the symposium incorporated valuable insights from industry experts, who shared real-world application scenarios and the specific requirements that industries expect from AI technologies.

In addition, the symposium featured two captivating challenges that engaged the community and encouraged their active participation in a follow-up event. The first challenge, presented by Paulo Shakarian from Arizona State University, centered around the independent evaluation of ChatGPT's performance on mathematical word problems. Shakarian proposed a benchmark dataset specifically designed for assessing the capabilities of chatbot systems in solving mathematical word problems in natural language. This challenge aimed to push the boundaries of AI in the realm of mathematical problem-solving.

The second challenge, by Maaïke de Boer and Roos Bakker from TNO, revolved around the dynamic ontology matching challenge. Recognizing the pressing challenges faced by the labor market, they proposed developing a novel ontology-matching approach for aligning ontologies related to labor market dynamics. This challenge addresses the labor market's friction between demand and supply, highlighting the potential of knowledge engineering and machine learning in offering innovative solutions.

By incorporating these challenges, the symposium anticipates fostering knowledge exchange and collaboration and provides a platform for researchers, practitioners, and students to tackle real-world problems and explore the potential of combining machine learning and knowledge engineering in practical applications.

3. Reflections in Keynotes

Edward A. Feigenbaum [6] remarked that most AI research and development is done in perception/recognition (such as statistical, data-oriented, and deep learning approaches), which is a highly competitive field. In addition, he pointed out that the cognition (such as reasoning and logic-based approaches) field has the potential to yield significant breakthroughs as it is mainly unsolved and less crowded. Feigenbaum also raised an intriguing question about the boundary between perception and cognition in AI, wondering if behaviors currently considered “cognitive” could become “perceptual.” He proposes exploring how much “thinking” is actually “recognizing” could be a promising research theme. Lastly, Feigenbaum suggests that young AI researchers should focus on less crowded areas, particularly investigating the boundary between perception and cognition.

Douglas B. Lenat [7] highlighted, in particular, the question of why current LLMs seem untrustworthy and brittle. Lenat elaborated on using a knowledge-based system as a source of truth to bias LLMs towards correctness and showcased the results of experiments on using LLMs, in this case GPT-3 [8], to generate CycL [9] “sentences.”

The possible utilization of LLMs in collaboration with reasoning systems has been stressed along the same lines in the opening symposium speech by Andreas Martin [10]. As depicted in Figure 1, Martin showcased the possible utilization of a probabilistic language model (LM) with instruction training [11], e.g., ChatGPT [12], that generates RDF(S) triples [13, 14, 15] from textual statements, with the potential to perform RDF(S) reasoning and constraint verification, as a new way to accomplish knowledge engineering.

The illustrated approach has been represented using a boxology [16, 17] that describes a hybrid intelligence [18] use case where text data (a prompt) is fed into a machine learning (ML) component to generate symbolic RDF(S) code. Subsequently, machine reasoning is performed as part of knowledge engineering (KE) to further infer explicit knowledge as RDF(s) triples.

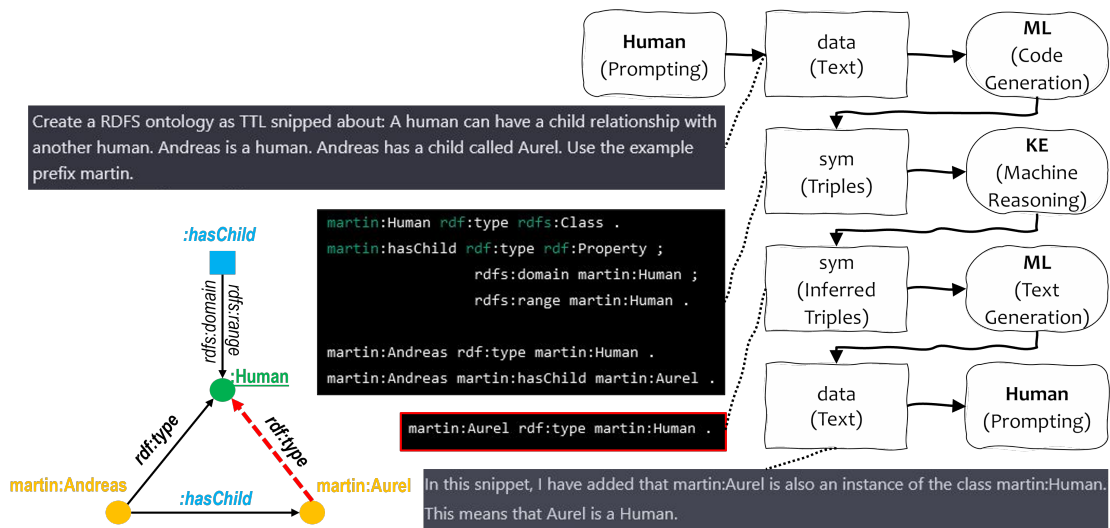


Figure 1: Boxology representation, ChatGPT prompt, RDF(S)/TTL output and visualization [10].

The resulting knowledge can then be transformed using an ML component for text generation and verbalization, converting it into natural language text data. This allows the gained inferences to be expressed in human-readable text. Figure 1 shows a prompt with common-sense knowledge on the top left as schema definition and instance description, which has been engineered by a human and then sent to ChatGPT. Below the prompt in Figure 1, the response of ChatGPT is presented. In this straightforward use case with this particular prompt given, the results of this not representative experiment were correct in all cases. However, further investigations in this field with more complex use cases are needed. Additional experiments on inferring types resulted in randomness, with the language model occasionally generating hallucinations by making up triples and rules.

As these LMs are probabilistic next-token predictors [19] and the requested reasoning seems to be simulated, it can be doubted whether reasoning at the level of RDF(S) can be achieved at all. It appears that the used LM is just trying to replicate and adapt RDF(S) triples that were already present in the training dataset, which were originally obtained from publicly accessible code repositories. Moreover, even in this simple, seemingly harmless example, it can be discussed that biases and stereotypes, e.g., about names and who is having children, through LLMs in particular [20], can be injected here. This also speaks for a possible inclusion of a general common-sense knowledge base verified ethically and under diversity aspects.

In conclusion, the experiment in Figure 1 demonstrates how machine learning and knowledge engineering can work together. The generated triples can be verified through constraint checking, ontology matching/alignment, RDF(S) reasoning, or human input in a human-in-the-loop setting [17].

4. Conclusion

The participants unanimously regarded the AAAI-MAKE symposium 2023 as an exceptionally successful, inspiring, and thought-provoking event, which effectively showcased the cutting-edge advancements and exemplified the immense potential and value derived from the integration of machine learning and knowledge engineering in the realm of AI research and practice. Moreover, the symposium served as a vibrant platform facilitating networking opportunities, nurturing meaningful interactions, and fostering fruitful contributions among participants from diverse backgrounds and domains.

As the symposium drew to a close, an engaging discussion transpired regarding a follow-up event, which could take the form of a subsequent symposium focusing on assessing approaches that combine knowledge engineering and machine learning and featuring captivating challenge presentations. The participants expressed their unwavering interest and enthusiasm for further collaboration and the ongoing exchange of ideas surrounding this crucial topic.

Overall, the AAAI-MAKE symposium 2023 left a lasting impression, igniting a collective commitment among the participants to sustain their collaborative efforts and propel advancements in the fusion of machine learning and knowledge engineering, ultimately charting new frontiers in AI.

Acknowledgments

Firstly, the AAAI-MAKE community would like to thank the two keynote speakers, Edward A. Feigenbaum and Douglas B. Lenat, for their exceptionally inspiring joint keynote. Further, thanks go to Maaïke de Boer (TNO), who delivered a short talk on behalf of AAAI-MAKE during the 2023 AAAI spring symposia plenary session.

We want to acknowledge the work of the organizing committee consisting of Reinhard Stolle, Doug Lenat, Hans-Georg Fill, Aurore Gerber, Knut Hinkelmann, Frank van Harmelen, and the symposium chair, Andreas Martin.

We would also like to thank the session chairs, Maaïke de Boer, Reinhard Stolle, Thomas Schmid, Emanuele Laurenzi, and Wilfrid Utz, and our various program committee members acting as reviewers. Finally, we would like to acknowledge Emanuele Laurenzi, who actively helped with the organization, and Charuta Pande, who compiled the proceedings.

References

- [1] A. Martin, AAAI-MAKE 2023: Challenges requiring the combination of machine learning and knowledge engineering, *AI Magazine* (2023). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/aaai.12094>. doi:10.1002/aaai.12094.
- [2] A. Martin, K. Hinkelmann, A. Gerber, D. Lenat, F. van Harmelen, P. Clark (Eds.), Proceedings of the AAAI 2019 Spring Symposium on Combining Machine Learning with Knowledge Engineering (AAAI-MAKE 2019), Stanford University, Palo Alto, California, USA, March 25-27, 2019, volume 2350 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019. URL: <https://ceur-ws.org/Vol-2350>.
- [3] A. Martin, K. Hinkelmann, H. Fill, A. Gerber, D. Lenat, R. Stolle, F. van Harmelen (Eds.), Proceedings of the AAAI 2020 Spring Symposium on Combining Machine Learning and Knowledge Engineering in Practice, AAAI-MAKE 2020, Palo Alto, CA, USA, March 23-25, 2020, Volume I, volume 2600 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2020. URL: <https://ceur-ws.org/Vol-2600>.
- [4] A. Martin, K. Hinkelmann, H. Fill, A. Gerber, D. Lenat, R. Stolle, F. van Harmelen (Eds.), Proceedings of the AAAI 2021 Spring Symposium on Combining Machine Learning and Knowledge Engineering (AAAI-MAKE 2021), Stanford University, Palo Alto, California, USA, March 22-24, 2021, volume 2846 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2021. URL: <https://ceur-ws.org/Vol-2846>.
- [5] A. Martin, K. Hinkelmann, H. Fill, A. Gerber, D. Lenat, R. Stolle, F. van Harmelen (Eds.), Proceedings of the AAAI 2022 Spring Symposium on Machine Learning and Knowledge Engineering for Hybrid Intelligence (AAAI-MAKE 2022), Stanford University, Palo Alto, California, USA, March 21-23, 2022, volume 3121 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2022. URL: <https://ceur-ws.org/Vol-3121>.
- [6] E. Feigenbaum, AAAI-MAKE 2023 - Keynote - Personal Perspectives on AI. In Zenodo [Keynote]. AAAI Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAAI-MAKE 2023), Hyatt Regency, San Francisco

- Airport, California, USA, 2023. URL: <https://zenodo.org/record/8015313>. doi:10.5281/zenodo.8015313.
- [7] D. Lenat, AAI-MAKE 2023 - Keynote - Personal Perspectives on AI. In Zenodo [Keynote]. AAI Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAI-MAKE 2023), Hyatt Regency, San Francisco Airport, California, USA, 2023. URL: <https://zenodo.org/record/8015329>. doi:10.5281/zenodo.8015329.
- [8] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language Models are Few-Shot Learners, in: H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, H. Lin (Eds.), *Advances in Neural Information Processing Systems*, volume 33, Curran Associates, Inc., 2020, pp. 1877–1901. URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf.
- [9] D. Lenat, R. V. Guha, CYC: A Midterm Report, *AI Magazine* 11 (1990) 32. URL: <https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/842>. doi:10.1609/aimag.v11i3.842.
- [10] A. Martin, AAI-MAKE 2023 - Opening Talk. In Zenodo [Keynote]. AAI Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAI-MAKE 2023), Hyatt Regency, San Francisco Airport, California, USA, 2023. URL: <https://zenodo.org/record/7996970>. doi:10.5281/zenodo.7996970.
- [11] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. F. Christiano, J. Leike, R. Lowe, Training language models to follow instructions with human feedback, in: S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, A. Oh (Eds.), *Advances in Neural Information Processing Systems*, volume 35, Curran Associates, Inc., 2022, pp. 27730–27744. URL: https://proceedings.neurips.cc/paper_files/paper/2022/file/b1efde53be364a73914f58805a001731-Paper-Conference.pdf.
- [12] J. Schulman, OpenAI, ChatGPT: Optimizing Language Models for Dialogue, 2022. URL: <https://web.archive.org/web/20221130180912/https://openai.com/blog/chatgpt/>.
- [13] W3C, 2014, RDF Schema 1.1, URL: <http://www.w3.org/TR/rdf-schema/>.
- [14] W3C, 2014, RDF 1.1 Turtle, URL: <http://www.w3.org/TR/turtle/>.
- [15] W3C, 2014, RDF 1.1 Concepts and Abstract Syntax, URL: <http://www.w3.org/TR/rdf11-concepts/>.
- [16] F. van Harmelen, A. ten Teije, A Boxology of Design Patterns for Hybrid Learning and Reasoning Systems, *Journal of Web Engineering* 18 (2019) 97–124. doi:10.13052/jwe1540-9589.18133.
- [17] H. F. Witschel, C. Pande, A. Martin, E. Laurenzi, K. Hinkelmann, Visualization of Patterns for Hybrid Learning and Reasoning with Human Involvement, Springer International Publishing, Cham, 2021, pp. 193–204. doi:10.1007/978-3-030-48332-6_13.
- [18] Z. Akata, D. Balliet, M. D. Rijke, F. Dignum, V. Dignum, G. Eiben, A. Fokkens, D. Grossi, K. Hindriks, H. Hoos, H. Hung, C. Jonker, C. Monz, M. Neerincx, F. Oliehoek, H. Prakken, S. Schlobach, L. V. D. Gaag, F. V. Harmelen, H. V. Hoof, B. V. Riemsdijk, A. V. Wynsberghe,

- R. Verbrugge, B. Verheij, P. Vossen, M. Welling, A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect with Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence, *Computer* 53 (2020). doi:10.1109/MC.2020.2996587.
- [19] S. Wolfram, *What is ChatGPT doing ... and why does it work?*, first edition ed., Wolfram Media, Inc, Champaign, Illinois, 2023.
- [20] E. M. Bender, T. Gebru, A. McMillan-Major, S. Shmitchell, On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?, in: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT '21*, Association for Computing Machinery, New York, NY, USA, 2021, p. 610–623. URL: <https://doi.org/10.1145/3442188.3445922>. doi:10.1145/3442188.3445922.