EDITOR: Amit Sheth, amit@sc.edu

DEPARTMENT: NEUROSYMBOLIC AI

Neurosymbolic Value-Inspired Artificial Intelligence (Why, What, and How)

Amit Sheth and Kaushik Roy, *University of South Carolina, Columbia, SC, 29208, USA*

The rapid progression of artificial intelligence (AI) systems, facilitated by the advent of large language models (LLMs), has resulted in their widespread application to provide human assistance across diverse industries. This trend has sparked significant discourse centered around the ever-increasing need for LLM-based AI systems to function among humans as a part of human society. Toward this end, neurosymbolic AI systems are attractive because of their potential to enable and interpretable interfaces for facilitating value-based decision making by leveraging explicit representations of shared values. In this article, we introduce substantial extensions to Kahneman's System 1 and System 2 framework and propose a neurosymbolic computational framework called value-inspired AI (VAI). It outlines the crucial components essential for the robust and practical implementation of VAI systems, representing and integrating various dimensions of human values. Finally, we further offer insights into the current progress made in this direction and outline potential future directions for the field.

Since the inception of artificial intelligence (AI) systems, a primary goal has been their seamless integration into human society, aiming to assist in demanding tasks, such as large-scale automation. Consequently, discussions about their responsible utilization, particularly as they gain advanced capabilities, have been integral to active and interdisciplinary academic discourse. Questions have arisen about the values embedded in these systems and, more broadly, how to ensure that their use benefits humankind. In recent times, the exceptional capabilities of large language models (LLMs) in AI have accelerated the widespread adoption of AI across diverse industries. However, this adoption has not been without profound social consequences for human users, giving rise to unforeseen social risks like biases and ethical concerns. In response to these risks, there is an urgent call to implement "controls" on LLMs and their outcomes.[1] For humans functioning within a society, the basis for providing such "controls" on their functioning is rooted in a set of shared values that enjoy broad consensus among the populace. These values span various dimensions, encompassing ethics, sociocultural norms, policies, regulations, laws, and other pertinent aspects.[2] Due to the complexity of such a value-based framework, decision-making behaviors in these situations are usually established by trained personnel in government positions with support from legislation on behalf of the wider population within society.[3] We make an important distinction between these carefully thought-out, expert-defined societal values for the synergistic functioning of humans within society, and preference patterns of a large collective of people, which may itself contain implicit notions of values but may also consequently contain population-wide biases. The latter does not readily clarify an unambiguous value-based stance that considers maintaining societal order and is therefore not suited for incorporation within value-inspired AI (VAI) systems without additional audits and checks by regulatory bodies (see "Why Value-Inspired Artificial Intelligence?").

1541-1672 © 2024 IEEE
Digital Object Identifier 10.1109/MIS.2023.3344353
Date of current version 7 February 2024.

AI-ASSISTED DRIVING MOTIVATING VAI SYSTEMS

The following scenario represents a case where a clearly defined value system is necessary for decision

WHY VALUE-INSPIRED ARTIFICIAL INTELLIGENCE?

As artificial intelligence (AI) systems increasingly take center stage in human assistance, we should expect the system to be aware of the values that a human operator would be aware of and adhere to. For example, an AI system used for driving assistance must be aware of the values of a human driver (e.g., values pertaining to driving rules, regulations, policies, and ethical behaviors laid out by the appropriate branches of governance) and abide by those values.

making and ensuring subsequent compliance with respect to acceptable values within human societies (e.g., sociocultural norms, policies, regulations, laws, and so on[4]). The example is extremely challenging and intended to motivate the imperative need for VAI frameworks. Consider a variant of the classic thought experiment known as the *trolley problem* in the AI-assisted driving setting that asks the following: Should you pull aside to divert your runaway vehicle so that it kills one person rather than five? Alternatively, what if a bicycle suddenly enters the lane? Should the vehicle swerve into oncoming traffic or hit the bicycle? Which choices are moral in these scenarios? One approach can be to decide based on the values that prioritize society, i.e., the fewest deaths, or a solution that values individual rights (such as the right not to be intentionally put in harm's way).[5] Ultimately, decisions are subject to a clearly defined and specified value system so that AI systems can be objectively and responsibly managed.

WHAT IS THE ROLE OF NEUROSYMBOLIC AI FOR VAI?

The Adequacy of Daniel Kahneman's Framework for VAI

As exemplified in the "AI-Assisted Driving Motivating VAI Systems" section, the mechanisms for value-inspired decision making in real-world scenarios can get quite complex and nuanced. Figure 1 shows a general architecture consisting of the essential components for human value-based decision making. We borrow from and extend Kahneman's System 1 and System 2 frameworks, which are the gold standard for formulating AI system objectives toward achieving human-like decision-making outcomes.[6] Figure 1 illustrates the two systems. System 2's functions involve the representation of deliberative "thought structures" (e.g., values, i.e., social-cultural-ethical norms, represented as graphs in the figure, and laws, rules, policies, and regulations, represented as a rule knowledge base in the figure) and the reasoning over them. System 1 involves faster and reflexive elements guided by familiar patterns, depicted by the use of neural networks in the figure. Neural networks are statistical methods that rely primarily on pattern recognition-based decision making. Traditionally, System 2 components lend themselves better to *symbolic* representations, and System 1 components lend themselves better to subsymbolic or *neural* representations (in the contemporary sense). Thus, it is natural to combine the two using a unified *neurosymbolic* framework.[7,8] However, as seen in the remainder of Figure 1, Kahneman's framework still needs extension because of its lack of specificity regarding the key requirements for a complete VAI architecture. We elucidate these specifics in the next sections.

Robust and Dynamic Knowledge Representation of Values

Current realizations of the System 2 part of Kahneman's framework lack definitions of the three primary facets of a value system required for the synergistic functioning of humans under a shared interpretation of values, especially in a society as diverse as humankind: facet 1 (f1) *unambiguous*, facet 2 (f2) *dynamic*, and facet 3 (f3) *rationalizable*. *Unambiguous* refers to a well-defined and precise interpretation of shared values, typically achieved through iterative procedures based on consensus among interested parties. *Dynamic* refers to the malleable nature of human values. Established human values do not change easily, however, they are subject to change depending on changing societal contexts, and in most cases, they do so organically, without much hindrance to the structures already in place. *Rationalizable*, which refers to the core axioms for the definitions in System 1, are mechanisms of change in System 2 and are grounded in principles of rational decision making, verifiable through some form of auditing or scrutiny (e.g., everybody can read the "constitution" and law books). We propose knowledge graphs as a viable symbolic representation of the values that meet the requirements for (f1), (f2), and (f3).

The semantic web community has dealt extensively with diverse and heterogeneous data sources and successfully integrated them into high-utility knowledge

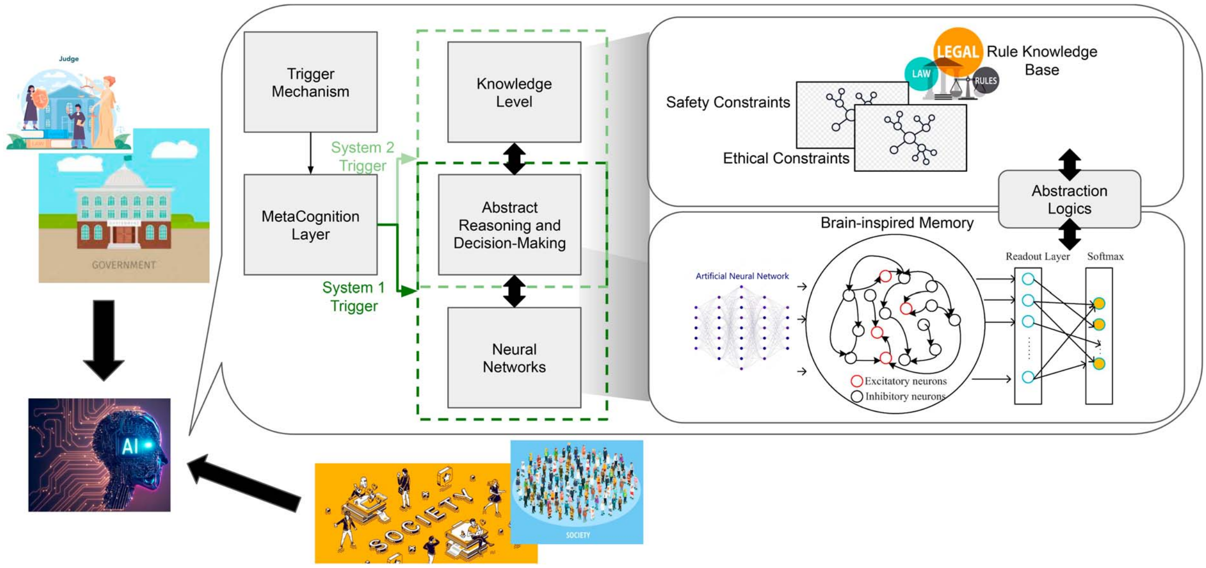

FIGURE 1. An illustration of the components of our proposed VAI framework. The components include a knowledge level (see the "Robust and Dynamic Knowledge Representation of Values" section); neural networks (see the "Brain-Inspired Memory Structures" section); abstraction logics, reasoning, and decision making (see the "Temporal Abstraction Logics" section); and metacognition layers with triggers (see the "Metacognition Layers and Triggers" section). We explain that a robust VAI method will need to integrate all these components using a neurosymbolic AI framework (see the "Neurosymbolic VAI: Putting It All Together" section).

representations, such as ontologies and knowledge graphs (e.g., Google Knowledge Graph, Wikipedia, and so on).[9] This ecosystem, centered around knowledge graphs, has proven its robustness in meeting information needs, specifically in regard to (f1), (f2), and (f3), across dimensions of quality, scale, and dynamic content changes. (F1) is achieved through processes for ensuring *ontological commitment*. (F2) is often a consequence of ensuring (f1), by which mechanisms for changes in the ontology design patterns and their practical implementation, both at the instance and ontology levels, are considered. (F3) is ensured through an established body of theoretical results on the aspects of rationality in machines, namely, soundness, completeness, verifiability, and decidability of knowledge graphs, and knowledge graph-based reasoners.[10] Today, knowledge graphs play a central role in various information processing and management tasks, including semantically enriched applications such as the web and other AI systems for search, browsing, recommendation, advertisement, and summarization in diverse domains.

Brain-Inspired Memory Structures

Transformer-based neural network architectures in LLMs have recently come to dominate the space of implementations for System 1 in Kahneman's framework. Established work on Brain-inspired cognitive architectures have made clear the distinctions among different types of memories based on their perceived roles. For example, declarative memory captures unchanging facts about the world, and episodic memory retains information about deviants from common schemas (e.g., birds that can't fly).[11] This plays a big part in several humans coexisting based on a common set or shared interpretation of values. One or more individuals may not like specific aspects of another's values, but this can be safely ignored as part of an episodic deviant as long as it is not a significant deal-breaker.

We first observe that systems such as LLMs are fundamentally large neural networks. Therefore, their loss surfaces, for most traditional choices of the loss function, are highly nonconvex, resulting in complex parameter-space dynamics, or *parametric-memory* dynamics. However, current training methods do not lead to models that adequately capture the dynamics. Specifically, the models do not possess *dynamic working memories*, and instead solve for a fixed point, i.e., a *static working memory*, which is invoked at inference time. The resulting inferences are expected to *generalize* to unseen test cases. We argue that to plausibly interact with the dynamic nature of the symbolic value representations, neural network architectures will need redesigns that capture a reasonable notion of *state dynamics*, thus making a distinction between episodic memories (a sequence of states) and generalizable memory (individual episode-agnostic generalizable patterns). Furthermore, because the neural network will need to interact with the symbolic layer, these changes will allow the network structures to have corresponding interfaces with appropriate parts of the symbolic layer. For example, generalizable memory interacts with generalized ontological constructs, and episodic memory interacts with instance-level changes or lack of conformance to the higher-level ontological constructs.

Temporal Abstraction Logics

Although Kahneman's framework defines individual types of systems based on their functions, it does not define how these systems will communicate with clarity. What might be a suitable modality, logic, or knowledge representation to enable such communication? To bridge this gap, we individually examine neural-network-based knowledge representations and more classical symbolic knowledge representations and propose a solution. Neural networks are adept at pattern recognition and consequently excel at capturing linguistic structure in the training data. Furthermore, The recent success of LLMs, particularly in benchmark tasks that test common sense, such as the Winograd challenge, shows evidence that *semantic* understanding does not always require a strictly structured propositional description. Rather, that semantic knowledge is deeply coupled with language patterns (e.g., linguistic and syntactic). At the same time, LLMs have performed embarrassingly poorly at tasks that require demonstration of several other aspects of common sense understanding (e.g., intuitive physics, planning, and causal sequence capture), thus showing the lack of a "full-bodied" understanding of the world. The "holy grail" of a full-bodied understanding of the world has been to adequately capture the full breadth of *relationships* between linguistic comprehension and semantic knowledge, one of the early endeavors of symbolic AI, e.g., WordNet (representing linguistic variations and word senses), ConceptNet (relationships among linguistic variations, word senses, and broader concepts and their properties), WikiData (relationships between concepts and entities such as people, places, things, and organizations), and so on.[12]

We therefore believe that the answer to a neurosymbolic interface for communication between System 1 and System 2 lies in a new kind of knowledge representation, which we refer to as *abstraction logics*—a mix of the classical propositional expressions and more distributional representations. Note that we are proposing a *blended* representation, which differs from previous work on neurosymbolic methods that either extract propositional representations from neural mechanisms or compress propositional information into continuous-valued vector spaces for consumption by neural networks. Moreover, existing neurosymbolic methods have concentrated primarily on a static portrayal of the world. In contrast, the abstraction logics developed in our framework must adeptly handle the dynamic nature of knowledge representations in System 2 and the dynamical model of neural networks proposed for System 1. This necessitates the incorporation of temporal aspects that capture the evolving nature of information. For instance, many value systems incorporate temporal dimensions (e.g., carbon emission goals for countries and corporations). We argue that the *relationships* expressed using these new kinds of *abstraction logics* will be at the heart of transitioning from machine comprehension of a string of sounds or letters to an internal representation of "meaning," which serves to enhance clarity in both AI and human interpretations of shared values.

Metacognition Layers and Triggers

Finally, Kahneman's framework does not specify when to invoke either system. Human decision making and the activities related to System 1 and System 2 are context specific and result in either reflexive decisions and actions based on prior experiences, or deliberative processes that involve slower thinking and reasoning. In the AI-assisted driving scenario discussed in the "AI-Assisted Driving Motivating VAI Systems" section, most of the decisions occur reflexively, guided by repetitive driving patterns. More intricate decisions, however, necessitate a deliberative process that considers the value structures for careful thought regarding subsequent steps and actions. Such decisions regarding when to choose between intricate versus reflexive action form aspects of *metacognition*, where a computational mechanism *triggers* either System 1 or System 2, or both, depending on the specific task.[13] This computational mechanism must be rooted in the fundamental interface between Systems 1 and 2, namely, the abstraction logics introduced in the "Temporal Abstraction Logics" section. Therefore, the practical implementation of a VAI architecture requires a module with functions geared toward implementing criteria that determine when System 1 and System 2 are invoked. This module is crucial to orchestrating the

(a) *Generative-AI based Frameworks Towards Value-Inspired AI*

Instruct-GPT Models utilize Reinforcement Learning from Human Feedback (RLHF) to align model responses with human-preferred ones, using a reward signal for desired behaviors [1].

Anthropic's Constitutional AI trains an AI assistant without specific human labels for harmful outputs, employing a "constitution" for oversight. Involves supervised and reinforcement learning phases, incorporating feedback through RL from AI Feedback (RLAIF) [2].

Social Science-Inspired Approaches draw inspiration from models like the helpful, honest, and harmless model and Schwartz's Theory of Basic Values. Represent social science dimensions in a value "vector" for seamless interaction with distributed vector-based representations in models like LLMs [3].

Civilizing AI Researchers at the AI Institute at the Univ. of South Carolina have coined the term Civilizing AI. First, they refer to challenges about AI-generated content identification as "eloquence." and unintended outcomes such as hallucinations as "adversity" and propose civilizing AI as Balancing AI's eloquence and adversarial tendencies [4].

REFERENCES

[1] Ouyang, Long, et al. ``Training language models to follow instructions with human feedback.'' Advances in Neural Information Processing Systems 35 (2022): 27730-27744
[2] Bai, Yuntao, et al. "Constitutional ai: Harmlessness from ai feedback." arXiv preprint arXiv:2212.08073 (2022).
[3] Yao, Jing, et al. "Value FULCRA: Mapping Large Language Models to the Multidimensional Spectrum of Basic Human Values." arXiv preprint arXiv:2311.10766 (2023).
[4] *AIISC*, aiisc.ai/Special.html. Accessed 07 Dec. 2023.

While these approaches exhibit potential, recent studies highlight their vulnerability to hallucinations and adversarial attacks, raising concerns about their controllability in achieving value-inspired AI objectives. Moreover, black-box large neural networks lack the foundational designs crucial for specifying effective controls, as outlined in Sections III-A,B,C,D. These sections emphasize the necessity of incorporating essential components to exert precise control over model behaviors for guiding them toward value-based outcomes.

(b) *Other Frameworks for Value-Inspired AI*

Humanity-Inspired AI Purohit et al. discuss a perspective on designing next-generation artificial intelligence (AI) systems driven by theoretical considerations. The focus is on envisioning AI systems with enhanced capabilities for meaningful human interaction. These systems aim to exhibit socially-adaptive behavior by incorporating personalization and an awareness of social context and intentionality. The proposed approach uses knowledge graphs, which combine general, common-sense, and domain-specific knowledge with socio-cultural values, norms, and individual cognitive models. The objective is to establish a foundation for developing AI systems that draw inspiration from humanity and its diverse aspects [5].

Societal Value-based Frameworks Work by Stanford's HAI group emphasizes the idea that today's AI models invariably carry underlying values, and, as a result, it's crucial to approach them with reflection and deliberation when determining which values to incorporate. By explicitly modeling societal values and leveraging advancements in natural language processing, there is the potential to represent a diverse range of concepts from the social and behavioral sciences. This approach offers a potent toolkit for influencing and shaping our online experiences [6].

REFERENCES

[5] Purohit, Hemant, Valerie L. Shalin, and Amit P. Sheth. "Knowledge graphs to empower humanity-inspired AI systems." *IEEE Internet Computing* 24.4 (2020): 48-54.
[6] Tuning Our Algorithmic Amplifiers: Encoding Societal Values into Social Media Ais." *Stanford HAI* (https://rb.gy/q1rrls).

Implementing next-gen AI systems driven by theoretical considerations poses challenges in translating theory into practical design. Creating socially-adaptive systems integrating personalization, social context, and intentionality requires addressing diverse individual preferences and societal dynamics. Implementing knowledge graphs that combine general, domain-specific knowledge, and socio-cultural values demands precise modeling. Explicitly modeling societal values in AI systems encounters challenges related to dynamic and evolving values.

(c) *Progress Towards Neurosymbolic Components (Section* III)

Independent efforts have been made toward progress components defined in Section III. For example, knowledge graph representations have expanded their scope to not only include declarative knowledge, but also a unified representation of both declarative and procedural knowledge with enhanced time-sensitive contexts (e.g., personalized vs. generalized contexts with dynamic causal effects, etc.) [7]. Control theory-inspired neural network designs, such as liquid time constant networks, do well in capturing parameter dynamics and have seen significant success in applications such as AI-assisted driving [8]. Established temporal logics, such as probabilistic circuits, and linear temporal logics, have been applied to specify logical constraints toward LLM safety control [9,10]. Finally, several ``cognitive architectures'' have been proposed to enable metacognitive functions, often guided by a sense of value, for orchestrating the functioning of individual components [11].

REFERENCES

[7] Sheth, Amit, et al. "Process knowledge-infused ai: Toward user-level explainability, interpretability, and safety." IEEE Internet Computing 26.5 (2022): 76-84.
[8] Hasani, Ramin, et al. "Liquid time-constant networks." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 35. No. 9. 2021.
[9] Ahmed, Kareem, Kai-Wei Chang, and Guy Van den Broeck. "A Pseudo-Semantic Loss for Autoregressive Models with Logical Constraints." Thirty-seventh Conference on Neural Information Processing Systems. 2023.
[10] Yang, Ziyi, et al. "Plug in the Safety Chip: Enforcing Temporal Constraints for LLM Agents." (2023).
[11] Sumers, Theodore, et al. "Cognitive architectures for language agents." arXiv preprint arXiv:2309.02427 (2023).

Despite substantial progress, a fully realized neurosymbolic method that consolidates components and ensuring synergistic functioning toward human values-guided functioning, is still a work in progress \cite{sheth2023neurosymbolic}. We explain the necessary components for a robust implementation of a Value-Inspired AI system and hope that our work will serve to foster impactful research along the directions mentioned in this paper.

FIGURE 2. Summary of our progress toward VAI. (a) Talks about generative AI-based conceptualizations of VAI frameworks by other researchers. (b) Talks about other frameworks that explicitly talk about integrating human values in society with AI systems. (c) Talks about progress toward the implementation of individual components mentioned in the "Robust and Dynamic Knowledge Representation of Values," "Brain-Inspired Memory Structures," "Temporal Abstraction Logics," and "Metacognition Layers and Triggers" sections.

interplay between reflexive and deliberative decision-making processes, facilitating the necessary dynamic and context-sensitive AI system responses.

Neurosymbolic VAI: Putting It All Together

As discussed in the "Robust and Dynamic Knowledge Representation of Values," "Brain-Inspired Memory Structures," "Temporal Abstraction Logics," and "Metacognition Layers and Triggers" sections, the synthesis of a comprehensive and capable neurosymbolic VAI architecture can be achieved by integrating the aforementioned components. Each component is equipped with specific implementations tailored to its unique functions: a dynamic knowledge-graph-centered network and reasoning mechanisms for robustly representing values and facilitating decision making based on these values (System 2), brain-inspired neural-network-based dynamical systems for the expressive encoding of various aspects of memory (System 1), temporal abstraction logics to facilitate the interaction between Systems 1 and 2, and metacognition layers and triggers that orchestrate the overall functioning of the AI system. This integrated architecture enables a computational framework capable of harmonizing diverse cognitive processes and enhancing the AI system's adaptability and responsiveness during operation with humans in human society under a shared value system.

HOW FAR ALONG WE ARE AND FUTURE DIRECTIONS

Figure 2 provides an overview of the progress that researchers have made thus far in incorporating human values within AI systems.

Generative-AI-Based Frameworks Toward VAI

Figure 2(a) describes the techniques used by three main classes of techniques for incorporation of human values, specifically incorporation by aligning model outputs with human preferences. Although these approaches exhibit potential, recent studies highlight their vulnerability to hallucinations and adversarial attacks, raising concerns about their controllability in achieving VAI. Moreover, large black-box neural networks lack the foundational designs crucial for specifying effective controls, as outlined in in the "Robust and Dynamic Knowledge Representation of Values," "Brain-Inspired Memory Structures," "Temporal Abstraction Logics," and "Metacognition Layers and Triggers" sections. These sections emphasize the necessity of incorporating essential components to exert precise control over model behaviors for guiding them toward value-based outcomes.

Other Frameworks for VAI

Figure 2(b) shows previous efforts that have characterized the complexity of human values within society. These works emphasize key challenges related to the dynamic and evolving nature of explicitly modeling societal values in AI systems and propose knowledge graphs and LLMs as candidates to handle such dynamics. Adapting LLMs, or general neural network processing techniques that incorporate the values represented in KGs, requires a computing framework for forming a clear understanding of value-based perspectives and considerations within the neural network's internal structures. These works have not concretely talked about such a framework. In this work, we provide a road map with concrete steps and specific implementation strategies for achieving VAI through a neurosymbolic computational method.

Progress Toward Neurosymbolic Components (the "What Is the Role of Neurosymbolic AI for VAI?" section)

Figure 2(c) shows substantial progress across all the components defined in the "What Is the Role of Neurosymbolic AI For VAI?" section. Despite advances in the individual components, a fully realized neurosymbolic method that consolidates components and ensures synergistic functioning toward human-values-guided functioning is still a work in progress.[14]

CONCLUDING REMARKS

In this article, we introduced VAI systems and expanded on Kahneman's Systems 1 and 2 framework by providing detailed outlines of the components necessary for robust implementation of VAI systems. Specifically, we identify existing implementation challenges and present a clear road map for integrating explicit models of societal values in knowledge graphs, paired with technical advances in neural methods, abstraction logics, and metacognition methods, within a neurosymbolic computational framework. We hope our work will inspire building on the considerable progress and further stimulate impactful research in the directions mentioned in this article.

ACKNOWLEDGMENT

This research was supported in part by National Science Foundation Award 2335967, "EAGER: Knowledge-Guided Neurosymbolic AI," with guardrails for safe virtual health assistants," (the opinions are those of authors and not the sponsor), and feedback from AI Institute of South Carolina colleagues.

REFERENCES

1. L. Weidinger et al., "Ethical and social risks of harm from language models," 2021, *arXiv:2112.04359*.
2. S. H. Schwartz and W. Bilsky, "Toward a universal psychological structure of human values," *J. Personality Social Psychol.*, vol. 53, no. 3, pp. 550–562, 1987, doi: 10.1037/0022-3514.53.3.550.
3. W. Leal Filho et al., "The role of governance in realising the transition towards sustainable societies," *J. Cleaner Prod.*, vol. 113, pp. 755–766, Feb. 2016, doi: 10.1016/j.jclepro.2015.11.060.
4. H. Purohit, V. L. Shalin, and A. P. Sheth, "Knowledge graphs to empower humanity-inspired AI systems," *IEEE Internet Comput.*, vol. 24, no. 4, pp. 48–54, Jul./Aug. 2020, doi: 10.1109/MIC.2020.3013683.
5. M. Geisslinger, F. Poszler, J. Betz, C. Lütge, and M. Lienkamp, "Autonomous driving ethics: From trolley problem to ethics of risk," *Philosophy Technol.*, vol. 34, no. 4, pp. 1033–1055, 2021, doi: 10.1007/s13347-021-00449-4.
6. D. Kahneman, *Thinking, Fast and Slow*. New York, NY, USA: Macmillan, 2011.
7. P. Hitzler and M. K. Sarker, *Neuro-Symbolic Artificial Intelligence: The State of the Art*. Amsterdam, The Netherlands: IOS Press, 2022.
8. A. Garcez and L. C. Lamb, "Neurosymbolic AI: The 3rd wave," *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 1–20, 2023, doi: 10.1007/s10462-023-10448-w.
9. A. Sheth, S. Padhee, and A. Gyrard, "Knowledge graphs and knowledge networks: The story in brief," *IEEE Internet Comput.*, vol. 23, no. 4, pp. 67–75, Jul./Aug. 2019, doi: 10.1109/MIC.2019.2928449.
10. H. J. ter Horst, "Combining RDF and part of OWL with rules: Semantics, decidability, complexity," in *Proc. 4th Int. Semant. Web Conf. (ISWC)*, Galway,

Ireland. Berlin, Germany: Springer-Verlag, 2005, pp. 668–684.
11. P. Langley, J. E. Laird, and S. Rogers, "Cognitive architectures: Research issues and challenges," *Cogn. Syst. Res.*, vol. 10, no. 2, pp. 141–160, Jun. 2009, doi: 10.1016/j.cogsys.2006.07.004.
12. J. Browning and Y. LeCun, "Language, common sense, and the Winograd schema challenge," *Artif. Intell.*, vol. 325, Dec. 2023, Art. no. 104031, doi: 10.1016/j.artint.2023.104031.
13. M. T. Cox, "Metacognition in computation: A selected research review," *Artif. Intell.*, vol. 169, no. 2, pp. 104–141, Dec. 2005, doi: 10.1016/j.artint.2005.10.009.
14. A. Sheth, K. Roy, and M. Gaur, "Neurosymbolic artificial intelligence (why, what, and how)," *IEEE Intell. Syst.*, vol. 38, no. 3, pp. 56–62, May/Jun. 2023, doi: 10.1109/MIS.2023.3268724.

AMIT SHETH is the founding director, NCR chair, and a professor of computer science and engineering at the University of South Carolina, Columbia, South Carolina, 29208, USA. Contact him at amit@sc.edu.

KAUSHIK ROY is a Ph.D. student in computer science at the AI Institute, University of South Carolina, Columbia, South Carolina, 29208, USA. Contact him at kaushikr@email.sc.edu.