

# Neuro-Symbolic AI in 2024: A Systematic Review

Brandon C. Colelough<sup>1,\*</sup>, William Regli<sup>2</sup>

<sup>1</sup>University of Maryland, College Park, 8125 Paint Branch Dr, College Park, MD 20742

## Abstract

**Background:** The field of Artificial Intelligence has undergone cyclical periods of growth and decline, known as AI summers and winters. Currently, we are in the third AI summer, characterized by significant advancements and commercialization, particularly in the integration of Symbolic AI and Sub-Symbolic AI, leading to the emergence of Neuro-Symbolic AI.

**Contributions:** (1) A definition of Meta-Cognition within Neuro-Symbolic AI. (2) A review of the key themes of the literature post the Neuro-Symbolic research explosion from 2020-2024. (3) Identification of the current gaps in the literature of Neuro-Symbolic AI

**Objective:** This paper provides a systematic literature review of Neuro-Symbolic AI projects within the 2020-24 AI landscape, highlighting key developments, methodologies, and applications. It aims to identify where quality efforts are focused in 2024 and pinpoint existing research gaps in the field.

**Methods:** The review followed the PRISMA methodology, utilizing databases such as IEEE Explore, Google Scholar, arXiv, ACM, and SpringerLink. The inclusion criteria targeted peer-reviewed papers published between 2020 and 2024. Papers were screened for relevance to Neuro-Symbolic AI, with further inclusion based on the availability of associated codebases to ensure reproducibility.

**Results:** From an initial pool of 1,428 papers, 167 met the inclusion criteria and were analyzed in detail. The majority of research efforts are concentrated in the areas of learning and inference (63%), logic and reasoning (35%), and knowledge representation (44%). Explainability and trustworthiness are less represented (28%), with Meta-Cognition being the least explored area (5%). The review identifies significant interdisciplinary opportunities, particularly in integrating explainability and trustworthiness with other research areas.

**Discussion:** The findings reveal a well-integrated body of work in learning and inference, logic and reasoning, and knowledge representation. However, there is a notable gap in research focused on explainability and trustworthiness, which is critical for the deployment of reliable AI systems. The sparse representation of Meta-Cognition highlights the need for further research to develop frameworks that enable AI systems to self-monitor, evaluate, and adjust their processes, enhancing autonomy and adaptability.

**Conclusion:** Neuro-Symbolic AI research has seen rapid growth since 2020, with concentrated efforts in learning and inference. Significant gaps remain in explainability, trustworthiness, and Meta-Cognition. Addressing these gaps through interdisciplinary research will be crucial for advancing the field towards more intelligent, reliable, and context-aware AI systems.

## Keywords

Neuro-Symbolic AI, Systematic Review, Learning and Inference, Knowledge Representation, Logic and Reasoning, Explainability and Trustworthiness, Meta-Cognition, PRISMA.

---

LNSAI 2024: First International Workshop on Logical Foundations of Neuro-Symbolic AI, August 05, 2024, Jeju, South Korea

\*Corresponding author.

✉ brandcol@umd.edu (B. C. Colelough); regli@umd.edu (W. Regli)

🌐 <https://brandoncolelough.com/> (B. C. Colelough); <https://www.cs.umd.edu/people/regli> (W. Regli)

🆔 0000-0001-8389-3403 (B. C. Colelough); 0000-0001-7116-9338 (W. Regli)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

## 1. Introduction

The field of Artificial Intelligence (AI) has experienced significant cyclical growth, known as AI summers and winters. At present, we as a community find ourselves in the third AI summer, marked by rapid scientific advances and commercialization, continuing the legacy of previous periods of AI excitement followed by setbacks [1]. A significant product of the third AI summer has been the integration of two prominent fields of AI; Symbolic AI and Sub-Symbolic AI, the fusion of which is known as Neuro-Symbolic AI. There is an ongoing debate about the necessity of Neuro-Symbolic AI [2], opponents arguing that common sense reasoning can be addressed through the use of big data [3] and proponents arguing that “You can’t get to the moon by climbing successively taller trees” [4]. For this systematic review, we take the stance that symbolic AI is essential and that Neuro-Symbolic AI represents the best way forward for the community hence, this paper provides a systematic literature review of prominent Neuro-Symbolic projects within the 2024 AI landscape, highlighting key developments, methodologies, and applications.

### 1.1. Symbolic AI

Symbolic AI is a “a sub-field of AI concerned with learning the internal symbolic representations of the world around it” where we can “translate some form of implicit human knowledge into a more formalized and declarative form based on rules and logic” [2]. Examples of some of the earliest AI systems that utilised symbolic representations include SHRDLU [5], ELIZA [6], DENDRAL [7] and MYCIN[8] and examples of some of the newest AI systems which heavily utilise symbolic processes include ConceptNet 5.5 [9] CYC [10] and Good Old Fashioned AI (GOFAI) planning systems [11] to name just a few.

### 1.2. Sub-Symbolic AI

By contrast, Sub-Symbolic AI are systems that “do not require rules or symbolic representations as inputs” and instead “learn implicit data representations on their own” [2]. Sub-Symbolic AI encompasses approaches such as machine learning, deep learning, and generative AI, which rely on algorithms to automatically extract patterns from raw data to discern relationships and make predictions based on learned representations. Examples of some of the earliest AI systems that utilised sub-symbolic representations include the Perceptron [12], Hopfield Networks [13] and the Backpropagation Algorithm [14] and examples of some of the newest sub-symbolic systems include famous projects such as the Generative Pre-trained Transformer (GPT) models [15], the YOLO family of Convolutional Neural Networks (CNN’S) [16] and the DALLE diffusion model transformer [17] to again just name a few.

### 1.3. Neuro-Symbolic AI

There is at present a debate within the AI community surrounding the need for Neuro-Symbolic AI [18]. Simply described, the argument for Neuro-Symbolic AI draws on Kahneman (2011) [19] concepts of System 1 and System 2 thinking whereby System 1 is fast, intuitive, and parallel, akin to the capabilities of deep learning, while System 2 is slow, deliberate, and sequential,

resembling symbolic reasoning and hence, Neuro-Symbolic AI aims to combine these two approaches to create systems that benefit from the strengths of both. We Adopt the definition provided by Garcez and Lamb (2023) [18]; Hence, Neuro-Symbolic AI is “**a composite AI framework that seeks to merge the domains of Symbolic AI and Neural Networks**” [or more broadly put, **Sub-Symbolic AI**] “**to create a superior hybrid AI model possessing reasoning capabilities**”. As this definition is quite broad, for the purpose of this systematic review, we will further define the sub-components of the Neuro-Symbolic AI taxonomy we believe to be most relevant to the current AI landscape within section 2.

## 2. Methodology

### 2.1. Taxonomy of Neuro-Symbolic AI

We identified five foundational research areas advancing the state of the art in Neuro-Symbolic AI. This taxonomy was synthesized from a review of six survey papers [20, 21, 22, 23, 24, 25] and four seminal books [2, 26, 27, 28]. These areas are:

1. **Knowledge Representation:** Integrating symbolic and neural representations and developing commonsense and domain-specific knowledge graphs [20, 26, 28].
2. **Learning and Inference:** Combining learning and reasoning processes through end-to-end differentiable reasoning and dynamic multi-source knowledge reasoning [21, 22, 2].
3. **Explainability and Trustworthiness:** Creating interpretable models and reasoning processes to ensure trust and reliability in Neuro-Symbolic systems [23, 24].
4. **Logic and Reasoning:** Integrating logic-based methods with neural networks, including logical and probabilistic reasoning, and the syntax and semantics of Neuro-Symbolic systems [23, 28].
5. **Meta-Cognition:** The system’s capacity to monitor, evaluate, and adjust its own reasoning and learning processes by integrating neural networks and symbolic representations.

The four above categories represent the core technical areas where current efforts are concentrated. Additionally, we define *Meta-Cognition* to address a gap in current taxonomies that fail to capture fields encompassing self-awareness, adaptive learning, reflective reasoning, self-regulation, and introspective monitoring.

### 2.2. Meta-Cognition

Meta-Cognition refers to the processes that involve thinking about one’s thinking, enabling self-awareness and self-regulation in cognitive tasks. Meta-Cognition is the controller that sits above the cognitive tasks of systems to direct energy effectively towards the correct system to handle a task. This higher-order cognition is crucial for tasks that require reflection, planning, and adaptation. Its importance lies in its ability to enhance learning, problem-solving, and

decision-making, making it a key focus in Neuro-Symbolic AI. Present research within Neuro-Symbolic AI does not yet effectively cover meta-cognition and neglecting Meta-Cognition in Neuro-Symbolic AI research limits system autonomy, adaptability, and reliability, hindering error correction and reducing trustworthiness in dynamic environments, making self-awareness and self-regulation essential for future advancements.

### 2.3. Literature Review Approach

We followed the PRISMA systematic review methodology to ensure a thorough and unbiased survey of the literature. Our search was conducted across five databases: IEEE Explore, Google Scholar, arXiv, ACM Digital Library, and SpringerLink Library, focusing on publications from 2020 to 2024. The keywords included "Neuro-Symbolic" combined with terms related to the foundational research areas. Only peer-reviewed articles, conference papers, and books in English were considered. From an initial broad search, we refined our selection to 392 candidate papers, which were further screened for quality, relevance, and availability of a public codebase, resulting in 167 papers that were included in our review. This streamlined process allowed us to focus on the most relevant and high-quality research, identifying key findings and open problems in Neuro-Symbolic AI. For a detailed description of our methodology, please refer to the full-length version of this survey paper, which will be uploaded to arXiv upon paper approval.<sup>1</sup>

## 3. Results

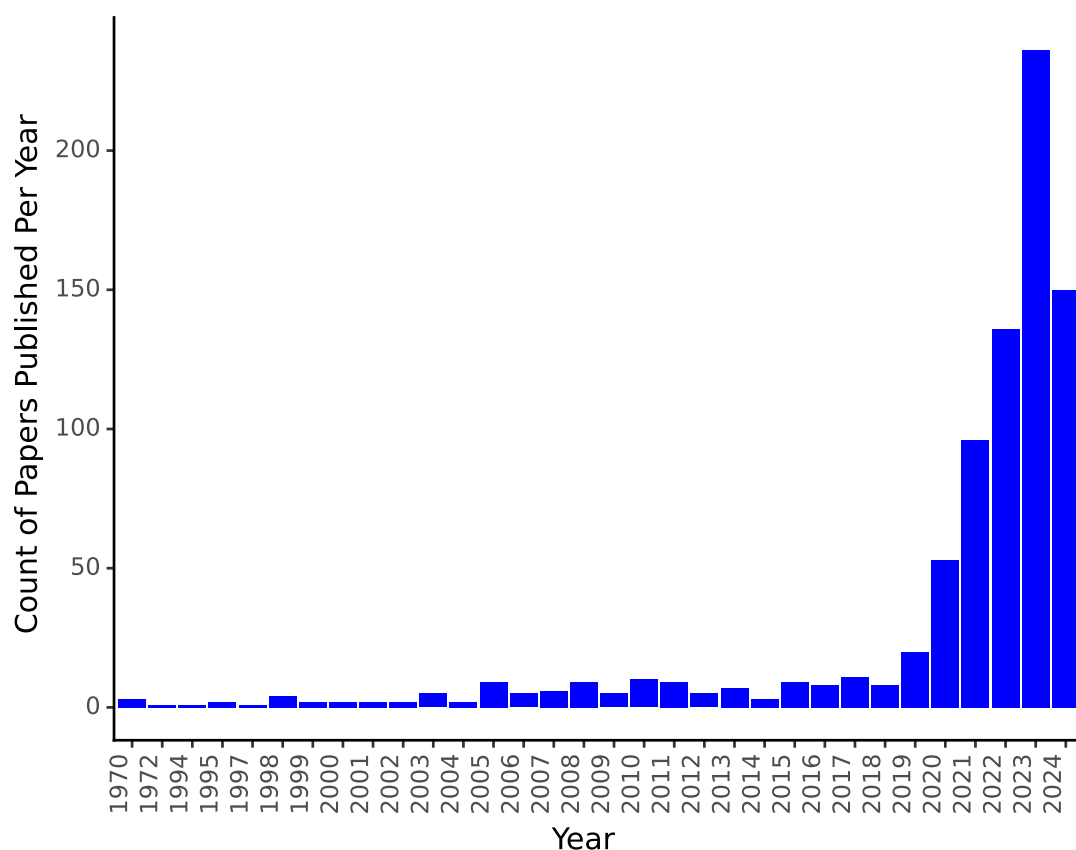
Database	Knowledge	Learning & Inference	Explainability & Trustworthiness	Logic & Reasoning	Meta-Cognition Cognition
IEEE	73	97	15	67	33
Google Scholar	56	126	7	129	3
ArXiv	17	54	7	55	3
ACM	10	46	5	12	17
Springer	152	170	65	162	47
Total(after screening)	308	493	99	425	103

**Table 1**

The search terms "neurosymbolic" AND each of the terms required for the 5 foundational research areas within neurosymbolic AI were queried through the 5 databases. The number of pieces of literature returned from each query is shown in the table above. Note also that only publications from 2020-2024 were considered

From the initial Google Scholar scraping, there was a total of 957 publications listed on Google Scholar alone from 1905 til the present. Figure 1 shows how research on Neuro-Symbolic AI is increasing exponentially starting in 2020, with notable increases in the years beginning from 2020 (53 publications), and peaking in 2023 (236 publications). Combining the Google Scholar literature from 2020 onwards with the literature from the four other databases queried for pieces of literature on Neuro-Symbolic AI from 2020 onwards, a total of 1,428 papers were extracted as

<sup>1</sup>Full manuscript will be uploaded to ARXIV upon paper approval.



**Figure 1:** Histogram of publications per year for Neuro-Symbolic AI. The data was obtained through Google Scholar scraping, reflecting significant growth from 2020

illustrated in table 1. From the total literature extracted, 45% ( $n = 641$ ) were removed as duplicate entries and 28% ( $n = 395$ ) were removed during title and abstract screening with 28% ( $n = 392$ ) held for further analysis. From the remaining 392 papers, the literature was further split based on code/model availability. 42% of the papers ( $n = 167$ ) had associated code-base repositories (e.g. GitHub, Huggingface etc.) and 58% ( $n = 225$ ) were further excluded from this literature review as a public code-base could not be found for the associated piece of literature (except for entries on Meta-Cognition as no code-bases could be found for literature associated with this category). The remaining 167 papers gathered were then read in detail and a further 9 papers were removed for not meeting the inclusion criteria leaving 158 included papers and 234 excluded papers. These 158 papers were then sub-categorised under the five main focal research areas and the intersection found therein. There were 44% ( $n=70$ ) entries in the Knowledge Representation category, 63% ( $n=99$ ) entries in the Learning and Inference category, 28% ( $n=44$ ) entries in the Explainability and Trustworthiness category, 35% ( $n=55$ ) entries in the Logic and Reasoning category, 5% ( $n=8$ ) entries in the Meta-Cognition category. The intersection of Knowledge Representation and Learning & Inference had 27% ( $n=43$ ) entries, the intersection of Knowledge

Representation and Explainability & Trustworthiness had 4% (n=5), the intersection of Learning & Inference and Logic & Reasoning had 11.48% (n=31), the intersection of Explainability & Trustworthiness and Logic & Reasoning had 3.33% (n=9). The intersection of any of the three categories had a range from 5 to 9 entries, except for the intersection of Explainability and Trustworthiness & Logic and Reasoning & Knowledge Representation, which had none. There was only one entry at the intersection of all 4 of the main research focal areas (excluding Meta-Cognition) which was AlphaGeometry from Google[29].

## 4. Discussion & Open Questions

To build upon the existing literature that has thoroughly summarized the Neuro-Symbolic landscape before 2020, we extend this discussion by analyzing the most influential projects in each sub-field of Neuro-Symbolic AI published since 2020. This section aims to highlight state-of-the-art (SOTA) technologies available to researchers, showcasing significant advancements and ongoing challenges.

### 4.1. Knowledge Representation

Research in Knowledge Representation has focused on advancing semantic grounding, representing complex relationships, and improving data efficacy. Development of commonsense knowledge bases and event-based representations [30, 31, 32] has advanced AI’s understanding of daily events, aiming to reduce error rates in text generation. These works are furthered through the exploration of minimal data requirements for commonsense knowledge in few-shot learning models[33], and the use of Neuro-Symbolic representations to enhance training efficiency and reduce costs [34]. Additionally, refinement of knowledge representation was demonstrated by predicting complex relationships and embedding techniques in knowledge graphs [35, 36], and the integration of personalized knowledge was demonstrated to ensure narrative consistency in storytelling agents [37]. NeuroQL, a domain-specific language for inter-subjective reasoning, captured complex and long-range relationships, demonstrating how Neuro-Symbolic approaches can ‘do more with less’, yielding significant savings in training time and environmental impact [38]. Open research questions remain around how Neuro-Symbolic AI can enhance the dynamic interpretation and manipulation of symbols, develop meta-cognitive abilities to monitor and adjust reasoning processes, and ensure transparent, explainable reasoning pathways for more human-like, adaptable, and robust knowledge representation.

### 4.2. Learning and Inference

Within Learning and Inference, research has focused on Neuro-Symbolic Integration for Enhanced Learning, Advanced Problem Solving and Decision Making, and Semantic Enhancement for Model Trustworthiness. Neuro-Symbolic integration for Enhanced Learning was demonstrated through the fusion of symbolic reasoning with neural learning mechanisms which adapted commonsense knowledge for few-shot settings and transformed observations into logical facts using Logical Neural Networks [39, 40]. Advanced Problem Solving and Decision

Making are highlighted by Plan-SOFAI [41] and the ZeroC architecture [42], which leverage Neuro-Symbolic methods to enhance AI planning and zero-shot concept recognition to integrate the fast and slow thinking models and the improve machine generalization. Semantic Enhancement and Model Trustworthiness were demonstrated by the introduction of a Pseudo-Semantic Loss for autoregressive models which integrated logic within the loss function [43] and neural networks utilising Logic Tensor Networks [44] which aim to boost logical consistency, reduce model toxicity, and enhance prediction accuracy by concentrating on relevant constraints. Open research questions remain in Neuro-Symbolic AI, including how to develop incremental learning that allows symbolic systems to evolve with new experiences, create context-aware inference mechanisms that adjust reasoning based on situational cues, achieve fine-grained explainability for complex inference chains, and explore meta-cognitive abilities enabling systems to monitor, evaluate, and optimize their learning processes in dynamic environments.

### 4.3. Explainability and Trustworthiness

Research centred on Explainability and Trustworthiness within Neuro-Symbolic AI has looked to advance Natural Language Processing (NLP) Techniques, Enhancing Logical Reasoning, and Refining Language Understanding and Summarization. Braid introduced a logical reasoner with probabilistic rules to tackle the brittle matching problem, merging symbolic and neural knowledge to enhance logical reasoning [45]. Similarly, Structure-Aware Abstractive Conversation improved summarization by incorporating discourse relations, action triples, and structured graphs for precise, context-rich summaries, advancing NLP techniques [46]. Semantic-level revisions identifying and correcting “confounders” in Neuro-Symbolic scenes have improved AI decision-making clarity, fostering trust and enhancing logical reasoning by making processes more understandable [47]. Evaluating AI’s humour comprehension with the New Yorker Cartoon Caption Contest underscored the need for nuanced understanding, refining language processing in complex cognitive tasks [48]. FactPEGASUS focuses on ensuring factuality in summarization by optimizing pre-training and fine-tuning methods, crucial for maintaining summary integrity and refining language understanding [49]. Complementing this, Neuro-Symbolic methods enhance explainable short answer grading through logical reasoning and cue detection, bridging AI capabilities with human-like responses and advancing NLP techniques [50]. Open research questions remain around how Neuro-Symbolic AI can adapt and evolve symbolic representations in real-time to maintain transparency, integrate meta-cognitive mechanisms for self-monitoring and adjustment of reasoning strategies, develop explainable NLP techniques for complex cognitive tasks, and ensure factual consistency in AI outputs while providing clear, detailed explanations of the underlying reasoning process.

### 4.4. Logic and Reasoning

From the research field of Neuro-Symbolic Logic and Reasoning, the research has gravitated largely towards the Integration of Logical Reasoning and Probabilistic Models, Commonsense Knowledge and Language Understanding and Enhanced Decision-Making. Logical Credal Networks [51] combines logical reasoning with probabilistic models to handle imprecise information, while DeepStochLog [52] enhances traditional logic programming with neural networks

for complex reasoning tasks. 2P-Kt [53] offers a comprehensive logic-based framework supporting various reasoning tasks and integrating symbolic and sub-symbolic AI. Research into Commonsense Knowledge and Language Understanding includes kogito, [32], which generates commonsense knowledge inferences from textual input to enhance AI adaptability, and LinkBERT [54], which improves language understanding and reasoning capabilities by incorporating document links, particularly in multi-hop reasoning tasks. For Enhanced Decision-Making, “Neuro-Symbolic Commonsense Social Reasoning” [55] integrates Neuro-Symbolic methods in autonomous systems, while LASER [49] combines neural networks’ flexibility with the precision of symbolic logic. “Getting from Generative AI to Trustworthy AI” [10] addresses LLMs’ limitations in trustworthiness and reasoning, proposing integration with symbolic AI systems for reliability. Open research questions remain around how Neuro-Symbolic AI can develop scalable frameworks that integrate traditional logic programming with neural networks for complex reasoning tasks, incorporate commonsense knowledge and advanced language understanding to enhance multi-hop reasoning capabilities, combine symbolic logic with neural networks to ensure reliable and trustworthy decision-making and integrate meta-cognitive abilities to enable self-monitoring and adjustment of reasoning strategies for clearer, more understandable explanations.

#### **4.5. Intersections of the above four research areas**

Much of the literature is cross-sectional between the four areas of research; Explainability & Trustworthiness, Knowledge Representation, Learning & Inference and Logic & Reasoning. AlphaGeometry [29], which is a Neuro-Symbolic system designed to solve Euclidean plane geometry problems at the Olympiad level, stands out as a prominent project that sits at the intersection of all four. AlphaGeometry’s ability to synthesise millions of theorems and proofs, using a neural language model trained on large-scale synthetic data to guide a symbolic deduction engine, makes it a groundbreaking example of how Neuro-Symbolic AI can achieve advanced problem-solving capabilities, bridging gaps across multiple domains of AI research. However, there is a distinct lack of integration with the explainability and trustworthiness fields within the unions of the other three research areas as the density of research intersection at the unions of explainability and trustworthiness and the other three research areas is relatively sparse, indicating a significant opportunity for further interdisciplinary work in Neuro-Symbolic AI.

#### **4.6. Meta-Cognition**

Recent advancements in this domain showcase Reinforcement learning (RL) to approximate Meta-Cognition, approaches to integrate cognitive architectures with LLMs to approximate meta-cognitive capabilities and the integration of many AI architectures and systems to demonstrate the Common Model of Cognition (CMC). The benefits of integrating symbolic features with RL algorithms were demonstrated through meta-reinforcement learning combined with logical program induction to improve financial trading strategies [56]. Enhancing general intelligence by fusing cognitive architectures with LLMs was investigated, creating embodied agents that leverage the strengths of both approaches [57, 58]. Improvements in experiential models were achieved by using LLMs to convert descriptive information into dense signals for instance-based



learning [59]. Adaptive conflict resolution in AI was enhanced by coupling cognitive reasoning with generative algorithms [60]. Robust AI systems were developed through modular, agency, and Neuro-Symbolic approaches to combine LLMs with cognitive architectures [61, 62]. These projects align with the CMC, integrating cognitive architectures like ACT-R, Soar, and Sigma to provide a unified framework for human cognition [63]. Enhancements in AI robustness and interoperability were achieved by integrating cognitive architectures with foundation models for cognitively guided few-shot learning [64]. Finally, combining generative networks with the CMC using a Neuro-Symbolic approach merged symbolic reasoning with neural networks to replicate human cognitive processes for powerful, explainable AI systems has been theorised but not yet realised[65]. Open research questions remain around how Neuro-Symbolic AI can integrate symbolic reasoning with meta-reinforcement learning for complex decision-making, fuse cognitive architectures with LLMs to develop meta-cognitive agents, leverage LLMs to enhance instance-based learning through meta-cognitive signals, create adaptive meta-cognitive frameworks for real-time conflict resolution, combine modular and agency approaches to build meta-cognitive AI systems aligned with the Common Model of Cognition, improve few-shot learning with cognitive architectures for meta-cognitive awareness, and develop Neuro-Symbolic generative networks that replicate human-like meta-cognitive processes.

## 5. Meta-Cognition in Neuro-Symbolic AI

Whilst the initial representation of Neuro-Symbolic AI as system 1 and system 2 level thinking is a useful tool to goal-orientate the field towards a common direction for the integration of neural and symbolic processes, the current adaptation of the human-level cognitive processing ability is too simplistic and does not yet capture the full systems-level breakdown of where the community should be investing effort to push the field forward. As Kahneman himself states, “the two systems do not really exist in the brain or anywhere else. ‘System 1 does X’ is a shortcut for ‘X occurs automatically.’ And ‘System 2 is mobilized to do Y’ is a shortcut for ‘arousal increases, pupils dilate, attention is focused, and activity Y is performed.’”. Human-level cognition manifests from a deeply complex and intricately layered yet densely connected box of systems of systems that work in unification and act with system-1 system-2 characteristics. The goal of the Neuro-Symbolic research domain is to “create a superior hybrid AI model possessing reasoning capabilities” and hence, to push the field further towards this goal we must seek to design and build systems that act with the same propensity as Kahnemans system-1 system-2 character setup through the implementation of more integrated systems of systems controlled through Meta-Cognition possessing the ability to act lazily when necessary and focused when required.

## 6. Conclusion

The field of Neuro-Symbolic AI has experienced a notable surge in research activity from 2020 onwards, reflecting the growing recognition of the importance of integrating symbolic and sub-symbolic approaches to enhance AI’s reasoning capabilities. The contribution from this systematic literature review is a well-grounded definition of Meta-Cognition within Neuro-Symbolic

AI, a review of the key themes of the literature post the Neuro-Symbolic research explosion from 2020-2024 and an identification of the current gaps in the literature of Neuro-Symbolic AI. We found that the majority of the research efforts in between 2020-24 were concentrated in the areas of learning and inference, with a significant portion also dedicated to logic and reasoning, as well as knowledge representation. These areas have seen substantial advancements, with innovative projects and methodologies pushing the boundaries of what AI systems can achieve in terms of understanding, reasoning, and generating human-like responses. However, our review also identifies several critical gaps in the current literature. Despite the substantial progress in learning and inference, there remains a relative sparseness of research focused on explainability and trustworthiness. This gap is particularly concerning given the increasing deployment of AI systems in real-world applications, where transparency and reliability are paramount. Moreover, the intersection of the four main research areas—learning and inference, logic and reasoning, knowledge representation, and explainability and trustworthiness—reveals a significant opportunity for interdisciplinary work. The density of studies that effectively combine these domains indicates the field is generally well integrated. The most underrepresented area in our review is Meta-Cognition. This emerging field, which involves systems' capacity to monitor, evaluate, and adjust their own reasoning and learning processes, holds great potential for advancing AI towards more autonomous and adaptable intelligence. The few existing studies in this domain suggest promising directions, but much more work is needed to develop robust frameworks and practical implementations of Meta-Cognitive architectures.

## References

- [1] H. A. Kautz, The third ai summer: Aaai robert s. engelmore memorial lecture, *AI Magazine* 43 (2022) 105–125. doi:10.1002/aaai.12036.
- [2] A. Dingli, Neuro-symbolic ai, 2023. Includes bibliographical references and index.
- [3] Y. LeCun, A path towards autonomous machine intelligence, *OpenReview Archive* (2022). URL: <https://openreview.net/forum?id=BZ5a1r-kVsf>, direct Upload.
- [4] G. Marcus, *Rebooting ai*, 2019.
- [5] T. Winograd, Procedures as a Representation for Data in a Computer Program for Understanding Natural Language, *AI Technical Report AITR-235*, 1971. URL: <http://hdl.handle.net/1721.1/7095>, accessed: 2024-05-13.
- [6] J. Weizenbaum, Eliza—a computer program for the study of natural language communication between man and machine, *Communications of the ACM* 9 (1966) 36–45. doi:10.1145/365153.365168.
- [7] R. K. Lindsay, B. G. Buchanan, E. A. Feigenbaum, J. Lederberg, *Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project*, McGraw-Hill, 1980.
- [8] W. van Melle, Mycin: a knowledge-based consultation program for infectious disease diagnosis, *International Journal of Man-Machine Studies* 10 (1978) 313–322. doi:10.1016/s0020-7373(78)80049-2.
- [9] R. Speer, J. Chin, C. Havasi, Conceptnet 5.5: An open multilingual graph of general knowledge (2016). doi:10.48550/ARXIV.1612.03975.
- [10] D. Lenat, G. Marcus, Getting from generative ai to trustworthy ai: What llms might learn from cyc, 2023. doi:10.48550/ARXIV.2308.04445.
- [11] S. Edelkamp, M. Helmert, R. Reffel, J. Bossek, International planning competition, in: *Proceedings of the Fourth International Conference on Automated Planning and Scheduling*, 2004, pp. 81–84.
- [12] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain., *Psychological Review* 65 (1958) 386–408. doi:10.1037/h0042519.
- [13] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities., *Proceedings of the National Academy of Sciences* 79 (1982) 2554–2558. doi:10.1073/pnas.79.8.2554.
- [14] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back-propagating errors, *Nature* 323 (1986) 533–536. doi:10.1038/323533a0.
- [15] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, 2017. doi:10.48550/ARXIV.1706.03762.

- [16] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, 2015. doi:10.48550/ARXIV.1506.02640.
- [17] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, I. Sutskever, Zero-shot text-to-image generation, 2021. doi:10.48550/ARXIV.2102.12092.
- [18] A. d. Garcez, L. C. Lamb, Neurosymbolic ai: the 3rd wave, *Artificial Intelligence Review* 56 (2023) 12387–12406. doi:10.1007/s10462-023-10448-w.
- [19] D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus and Giroux, 2011.
- [20] W. Gibaut, L. Pereira, F. Grassiotto, A. Osorio, E. Gadioli, A. Munoz, S. Gomes, C. d. Santos, Neurosymbolic ai and its taxonomy: a survey (2023). doi:10.48550/ARXIV.2305.08876.
- [21] D. Yu, B. Yang, D. Liu, H. Wang, S. Pan, A survey on neural-symbolic learning systems, 2021. doi:10.48550/ARXIV.2111.08164.
- [22] Z. Wan, C.-K. Liu, H. Yang, C. Li, H. You, Y. Fu, C. Wan, T. Krishna, Y. Lin, A. Raychowdhury, Towards cognitive ai systems: a survey and prospective on neuro-symbolic ai, 2024. doi:10.48550/ARXIV.2401.01040.
- [23] G. Marra, S. Dumančić, R. Manhaeve, L. De Raedt, From statistical relational to neurosymbolic artificial intelligence: A survey, *Artificial Intelligence* 328 (2024) 104062. doi:10.1016/j.artint.2023.104062.
- [24] C. Michel-Deletie, M. K. Sarker, Neuro-symbolic methods for trustworthy ai: a systematic review, *Neurosymbolic Artificial Intelligence* (2024). Tracking #: 726-1708, Review Assignment Stage.
- [25] D. Bouneffouf, C. C. Aggarwal, Survey on applications of neurosymbolic artificial intelligence, 2022. doi:10.48550/ARXIV.2209.12618.
- [26] P. Hitzler, M. K. Sarker, A. Eberhart (Eds.), *Compendium of Neurosymbolic Artificial Intelligence*, volume 369 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2023. URL: <https://doi.org/10.3233/FAIA369>. doi:10.3233/FAIA369.
- [27] P. Hitzler, M. K. Sarker, Neuro-symbolic artificial intelligence: The state of the art, in: *Neuro-Symbolic Artificial Intelligence*, 2021. URL: <https://api.semanticscholar.org/CorpusID:245698629>.
- [28] P. Shakarian, C. Baral, G. I. Simari, B. Xi, L. Pokala, *Neuro Symbolic Reasoning and Learning*, SpringerBriefs in Computer Science, Springer, 2023.
- [29] T. H. Trinh, Y. Wu, Q. V. Le, H. He, T. Luong, Solving olympiad geometry without human demonstrations, *Nature* 625 (2024) 476–482. URL: <https://github.com/google-deepmind/alphageometry>. doi:10.1038/s41586-023-06747-5.
- [30] N. Mostafazadeh, A. Kalyanpur, L. Moon, D. Buchanan, L. Berkowitz, O. Biran, J. Chu-Carroll, Glucose: Generalized and contextualized story explanations, 2020. URL: <https://github.com/ElementalCognition/glucose>. doi:10.48550/ARXIV.2009.07758.
- [31] J. D. Hwang, C. Bhagavatula, R. Le Bras, J. Da, K. Sakaguchi, A. Bosselut, Y. Choi, (comet-) atomic 2020: On symbolic and neural commonsense knowledge graphs, *Proceedings of the AAAI Conference on Artificial Intelligence* 35 (2021) 6384–6392. URL: <https://github.com/allenai/comet-atomic-2020>. doi:10.1609/aaai.v35i17.16792.
- [32] M. Ismayilzada, A. Bosselut, kogito: A commonsense knowledge inference toolkit, 2022. URL: <https://github.com/epfl-nlp/kogito>. doi:10.48550/ARXIV.2211.08451.
- [33] D. N. Ribeiro, K. Forbus, Combining analogy with language models for knowledge extraction, in: *3rd Conference on Automated Knowledge Base Construction*, 2021. URL: <https://github.com/dnr2/analogue-ke>. doi:10.59350/mk9dg-4r198.
- [34] K. Ahmed, S. Teso, K.-W. Chang, G. V. den Broeck, A. Vergari, Semantic probabilistic layers for neuro-symbolic learning, *arXiv preprint arXiv:2206.00426* (2022). URL: <https://github.com/thuwzy/Neural-Symbolic-and-Probabilistic-Logic-Papers>. doi:10.3233/faia230153, submitted 1 June, 2022; originally announced June 2022.
- [35] Z. Chen, G. Weiss, E. Mitchell, A. Celikyilmaz, A. Bosselut, Reckoning: Reasoning through dynamic knowledge encoding, 2023. URL: <https://github.com/eric11eca/reckoning-metaskg>. doi:10.48550/ARXIV.2305.06349.
- [36] A. Perevalov, D. Diefenbach, R. Usbeck, A. Both, Qald-9-plus: A multilingual dataset for question answering over dbpedia and wikidata translated by native speakers, 2022. URL: [https://github.com/perevalov/qald\\_9\\_plus](https://github.com/perevalov/qald_9_plus). doi:10.48550/ARXIV.2202.00120.
- [37] S. Gao, B. Borges, S. Oh, D. Bayazit, S. Kanno, H. Wakaki, Y. Mitsufuji, A. Bosselut, Peacock: Persona commonsense knowledge for consistent and engaging narratives, 2023. URL: <https://github.com/silin159/peacock>. doi:10.48550/ARXIV.2305.02364.
- [38] N. Papoulias, Neuroql: A neuro-symbolic language and dataset for inter-subjective reasoning, *arXiv preprint arXiv:2303.07146* (2023). URL: <https://github.com/orgdlabs/neuroQL>. doi:10.1109/cict56698.2022.9997814, submitted 13 March, 2023; originally announced March 2023.
- [39] J. Da, R. L. Bras, X. Lu, Y. Choi, A. Bosselut, Analyzing commonsense emergence in few-shot knowledge models, 2021. URL: <https://github.com/allenai/few-shot-comet>. doi:10.48550/ARXIV.2101.00297.
- [40] D. Kimura, M. Ono, S. Chaudhury, R. Kohita, A. Wachi, D. J. Agravante, M. Tatsubori, A. Munawar, A. Gray, Neuro-symbolic reinforcement learning with first-order logic, 2021. URL: <https://github.com/cav-research-lab/safe-reinforcement-learning-using-symbolic-logical-programming-for-autonomous-highway-driving>, <https://ibm.github.io/neuro-symbolic-ai/toolkit/>. doi:10.48550/ARXIV.2110.10963.
- [41] F. Fabiano, V. Pallagani, M. B. Ganapini, L. Hoesch, A. Loreggia, K. Murugesan, F. Rossi, B. Srivastava, Plan-SOFAI: A neuro-symbolic planning architecture, in: *Neuro-Symbolic Learning and Reasoning in the era of Large Language Models*, 2023. URL: <https://github.com/FrancescoFabiano/SOFAI-x-Planning>. arXiv: <https://openreview.net/forum?id=ORAhay0H4x>.
- [42] T. Wu, M. Tjandrasuwita, Z. Wu, X. Yang, K. Liu, R. Sosić, J. Leskovec, Zeroc: A neuro-symbolic model for zero-shot concept recognition and acquisition at inference time, *arXiv preprint arXiv:2206.15049* (2022). URL: <https://github.com/snap-stanford/zeroc>. doi:10.1109/tmm.2023.3318300, submitted 11 October, 2022; v1 submitted 30 June, 2022; originally announced

June 2022.

- [43] K. Ahmed, K.-W. Chang, G. V. d. Broeck, A pseudo-semantic loss for autoregressive models with logical constraints, 2023. URL: <https://github.com/UCLA-StarAI/PseudoSL>. doi:10.48550/ARXIV.2312.03905.
- [44] S. Badreddine, A. d'Avila Garcez, L. Serafini, M. Spranger, Logic tensor networks, *Artificial Intelligence* 303 (2022) 103649. URL: <https://github.com/logictensornetworks/logictensornetworks>. doi:10.1016/j.artint.2021.103649.
- [45] A. Kalyanpur, T. Breloff, D. Ferrucci, Braid: Weaving symbolic and neural knowledge into coherent logical explanations, 2020. URL: <https://github.com/chanind/tensor-theorem-prover>. doi:10.48550/ARXIV.2011.13354.
- [46] J. Chen, D. Yang, Structure-aware abstractive conversation summarization via discourse and action graphs, 2021. URL: <https://github.com/GT-SALT/Structure-Aware-BART>. doi:10.48550/ARXIV.2104.08400.
- [47] W. Stammer, P. Schramowski, K. Kersting, Right for the right concept: Revising neuro-symbolic concepts by interacting with their explanations, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 3618–3628. URL: <https://github.com/ml-research/NeSyXIL>. doi:10.1109/CVPR46437.2021.00362.
- [48] J. Hessel, A. Marasović, J. D. Hwang, L. Lee, J. Da, R. Zellers, R. Mankoff, Y. Choi, Do androids laugh at electric sheep? humor "understanding" benchmarks from the new yorker caption contest (2022). URL: [https://github.com/jmhessel/caption\\_contest\\_corpus](https://github.com/jmhessel/caption_contest_corpus). doi:10.48550/ARXIV.2209.06293.
- [49] J. Huang, Z. Li, M. Naik, S.-N. Lim, Laser: A neuro-symbolic framework for learning spatial-temporal scene graphs with weak supervision, arXiv preprint arXiv:2304.07647 (2023). URL: <https://github.com/isLinXu/paper-list>. doi:10.23919/eusipco58844.2023.10289877, submitted 22 November, 2023; v1 submitted 15 April, 2023; originally announced April 2023.
- [50] F. Künnecke, A. Filighera, C. Leong, T. Steuer, Enhancing multi-domain automatic short answer grading through an explainable neuro-symbolic pipeline, arXiv preprint arXiv:2403.01811 (2024). URL: <https://github.com/iwangjian/Paper-Reading-ConvAI>. doi:10.18653/v1/2023.bea-1.29, submitted 19 March, 2024; v1 submitted 4 March, 2024; originally announced March 2024.
- [51] H. Qian, R. Marinescu, A. Gray, D. Bhattacharjya, F. Barahona, T. Gao, R. Riegel, P. Sahu, Logical credal networks, 2021. URL: <https://github.com/radam2275/crema>. doi:10.48550/ARXIV.2109.12240.
- [52] T. Winters, G. Marra, R. Manhaeve, L. De Raedt, Deepstochlog: Neural stochastic logic programming, 2021. URL: <https://github.com/ML-KULeuven/deepstochlog>. doi:10.48550/ARXIV.2106.12574.
- [53] G. Ciatto, R. Calegari, A. Omicini, 2p-kt: A logic-based ecosystem for symbolic ai, *SoftwareX* 16 (2021) 100817. URL: <https://github.com/tuProlog/2p-kt>. doi:10.1016/j.softx.2021.100817.
- [54] M. Yasunaga, J. Leskovec, P. Liang, Linkbert: Pretraining language models with document links, 2022. URL: <https://github.com/michiyasunaga/LinkBERT>. doi:10.48550/ARXIV.2203.15827.
- [55] D. Chanin, A. Hunter, Neuro-symbolic commonsense social reasoning, arXiv preprint arXiv:2303.08264 (2023). URL: <https://github.com/chanind/amr-logic-converter>, submitted 14 March, 2023; originally announced March 2023.
- [56] S. I. Harini, G. Shroff, A. Srinivasan, P. Faldu, L. Vig, Neuro-symbolic meta reinforcement learning for trading, arXiv preprint arXiv:2302.08996 (2023). URL: <https://github.com/IBM/neuro-symbolic-ai>. doi:10.18653/v1/2021.emnlp-main.283, submitted 15 January, 2023; originally announced February 2023.
- [57] H. Joshi, V. Ustun, Augmenting cognitive architectures with large language models, *Proceedings of the AAAI Symposium Series 2* (2024) 281–285. URL: <https://github.com/ysymyth/awesome-language-agents>. doi:10.1609/aaaiss.v2i1.27689.
- [58] Y. Liu, Y. Liu, C. Shen, Combining minds and machines: Investigating the fusion of cognitive architectures and generative models for general embodied intelligence, *Proceedings of the AAAI Symposium Series 2* (2024) 307–314. URL: <https://github.com/Yangyi-Chen/Multimodal-AND-Large-Language-Models>. doi:10.1609/aaaiss.v2i1.27693.
- [59] C. McDonald, T. Malloy, T. N. Nguyen, C. Gonzalez, Exploring the path from instructions to rewards with large language models in instance-based learning, *Proceedings of the AAAI Symposium Series 2* (2024) 334–339. URL: [https://github.com/datainsightat/introduction\\_llm](https://github.com/datainsightat/introduction_llm). doi:10.1609/aaaiss.v2i1.27697.
- [60] A. Raja, A. Leshchenko, J. Kim, Leveraging conflict to bridge cognitive reasoning and generative algorithms, *Proceedings of the AAAI Symposium Series 2* (2024) 391–395. URL: <https://github.com/cmhungsteve/Awesome-Transformer-Attention>. doi:10.1609/aaaiss.v2i1.27705.
- [61] O. J. Romero, J. Zimmerman, A. Steinfeld, A. Tomic, Synergistic integration of large language models and cognitive architectures for robust ai: An exploratory analysis, *Proceedings of the AAAI Symposium Series 2* (2024) 396–405. URL: <https://github.com/ysymyth/awesome-language-agents>. doi:10.1609/aaaiss.v2i1.27706.
- [62] T. R. Sumers, S. Yao, K. Narasimhan, T. L. Griffiths, Cognitive architectures for language agents, 2023. URL: <https://github.com/ysymyth/awesome-language-agents>. doi:10.48550/ARXIV.2309.02427.
- [63] J. E. Laird, C. Lebiere, P. S. Rosenbloom, A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics, *AI Magazine* 38 (2017) 13–26. doi:10.1609/aimag.v38i4.2744.
- [64] R. H. Thomson, N. D. Bastian, Integrating cognitive architectures with foundation models: Cognitively-guided few-shot learning to support trusted artificial intelligence, *Proceedings of the AAAI Symposium Series 2* (2024) 409–414. doi:10.1609/aaaiss.v2i1.27708.
- [65] R. L. West, S. Eckler, B. Conway-Smith, N. Turcas, E. Tomkins-Flanagan, M. A. Kelly, Bridging generative networks with the common model of cognition, *Proceedings of the AAAI Symposium Series 2* (2024) 415–421.