

Classificação de Trabalhos em Inteligência Artificial Neuro-Simbólica a partir das Taxonomias de Kautz e de Bader & Hitzler

João P. Lepinsk¹, Thomas K. Derkascz¹, Rafael C. Cardoso²,
André P. Borges¹, Gleifer V. Alves¹

¹Universidade Tecnológica Federal do Paraná (UTFPR)
Ponta Grossa, PR, Brasil

²University of Aberdeen
Aberdeen, United Kingdom

{lepinsk,thomaskrevey}@alunos.utfpr.edu.br, rafael.cardoso@abdn.ac.uk

{apborges,gleifer}@utfpr.edu.br

Abstract. *Neuro-Symbolic Artificial Intelligence (NeSy AI) aims to integrate neural learning with the explainability of symbolic reasoning, combining performance and interpretability. However, the diversity of architectures makes comparisons and the construction of a clear overview of the field challenging. This work analyzes ten neuro-symbolic applications, selected through a systematic literature review, and classifies them according to two models: the taxonomy of Bader & Hitzler (2005), which evaluates the interrelation between components, type of symbolic language, and application purpose; and the framework proposed by Kautz (2021), which organizes systems into six integration architectures. The results show a predominance of hybrid approaches, frequent use of propositional languages, and highlight the Neuro(Symbolic) design. The analysis provides a clearer view of the field by revealing trends and gaps, as well as guiding future research and classifications of neuro-symbolic systems.*

Resumo. *A Inteligência Artificial Neuro-Simbólica (NeSy AI) busca integrar o aprendizado neural com a explicabilidade do raciocínio simbólico, aliando desempenho e interpretabilidade. Contudo, a diversidade de arquiteturas dificulta comparações e a construção de um panorama claro da área. Este trabalho analisa dez aplicações neuro-simbólicas, selecionadas por mapeamento sistemático da literatura, e as classifica segundo dois modelos: a taxonomia de Bader & Hitzler (2005), que avalia inter-relação entre componentes, tipo de linguagem simbólica e finalidade da aplicação; e a proposta de Kautz (2021), que organiza sistemas em seis arquiteturas de integração. Os resultados mostram a predominância de abordagens híbridas, o uso frequente de linguagens proposicionais e o destaque para o design Neuro(Symbolic). A análise oferece uma visão mais clara da área ao revelar tendências e lacunas, além de orientar pesquisas e classificações futuras de sistemas neuro-simbólicos.*

1. Introdução

A Inteligência Artificial Neuro-Simbólica (*Neuro-Symbolic Artificial Intelligence — NeSy AI*) é uma vertente da IA que busca unir o melhor dos paradigmas neural e simbólico na

construção de sistemas inteligentes [Saker et al. 2021]. A integração do processamento neural com regras simbólicas visa oferecer sistemas mais flexíveis, adaptativos e robustos, conforme aponta [Hagos et al. 2024].

No que diz respeito à parte simbólica, [Saker et al. 2021] a define como baseada na manipulação explícita de símbolos, utilizando lógica formal para realizar deduções e inferências. Esses sistemas, segundo [Dennis et al. 2023], apresentam características de transparência e explicabilidade.

Por outro lado, a parte neural, também denominada conexionista [Saker et al. 2021], é normalmente associada a sistemas considerados como “caixa-preta”, devido à dificuldade em fornecer explicações claras de seu funcionamento interno. Ainda assim, [Dennis et al. 2023] destaca que esses sistemas são rápidos e aptos a lidar com grandes volumes de dados.

Considerando a opacidade inerente aos sistemas conexionistas, torna-se relevante analisar e classificar as abordagens que propõem a integração dos aspectos simbólicos e neurais. Nesse sentido, [Bader and Hitzler 2005] propuseram uma classificação com o objetivo de organizar um campo marcado pela diversidade e dificuldade de comparação entre propostas. Sua taxonomia é baseada em três eixos — integração neuro-simbólica, finalidade do sistema e linguagem simbólica adotada — que, combinados, totalizam oito dimensões.

Complementarmente, [Kautz 2021] apresentou um novo esquema de classificação, composto por seis designs distintos, cada um descrevendo uma forma específica de integração entre componentes simbólicos e neurais. Dentre esses, destaca-se a arquitetura Neuro[Symbolic], apontada pelo autor como uma proposta ainda conceitual, sem implementações práticas conhecidas até então.

Este trabalho parte da compreensão dessas classificações e da análise do funcionamento dos sistemas neuro-simbólicos para realizar a categorização de dez artigos científicos com propostas relacionadas à área. A análise permitiu identificar tendências e padrões recorrentes nas soluções estudadas, contribuindo para uma visão mais estruturada e aprofundada do campo da NeSy AI.

O artigo está organizado da seguinte forma: na Seção 2, são apresentados os trabalhos relacionados; a Seção 3 descreve a metodologia adotada para a seleção e classificação dos artigos; na Seção 4, são expostos os resultados do mapeamento sistemático e a aplicação das taxonomias de Bader & Hitzler (2005) e Kautz (2021) aos textos selecionados; por fim, a Seção 5 discute os achados e aponta possibilidades para futuras pesquisas.

2. Trabalhos Relacionados

A área da NeSy AI tem sido alvo de diversos estudos que buscam não apenas o avanço tecnológico, mas também a organização e categorização das abordagens existentes. Nos últimos anos, surgiram revisões e propostas de taxonomias que visam sistematizar o entendimento sobre as arquiteturas híbridas e suas aplicações.

Dentre essas iniciativas, destaca-se o trabalho de Bekkum [van Bekkum et al. 2021]. Nesse estudo, os autores propõem uma taxonomia modular para sistemas híbridos, com foco na interação entre componentes simbólicos e conexionistas. Além disso, apresentam padrões recorrentes (*design patterns*) que

descrevem como essas arquiteturas podem ser construídas de forma reutilizável e clara, oferecendo exemplos concretos em diferentes domínios. O trabalho enfatiza a necessidade de modularidade e transparência nas soluções em NeSy AI, propondo um framework que permite comparar e projetar novos sistemas híbridos a partir de componentes bem definidos.

Outra contribuição relevante é o estudo de Sarker [Saker et al. 2021], que oferece uma revisão abrangente das tendências mais recentes em NeSy AI. O artigo propõe uma visão geral da evolução das arquiteturas, categorizando-as conforme seu grau de integração simbólica e neural, utilizando as taxonomias clássicas, como a de Bader & Hitzler [Bader and Hitzler 2005] e a proposta por Kautz [Kautz 2021], o trabalho também apresenta como diferentes soluções híbridas têm sido aplicadas, especialmente no que tange à explicabilidade e robustez dos sistemas.

Com foco mais específico na área de Processamento de Linguagem Natural (PLN), o trabalho de Hamilton [Hamilton et al. 2024], realiza um sistemático baseada diretamente na taxonomia de Kautz [Kautz 2021]. O estudo classifica as abordagens analisadas, permitindo uma visão clara das tendências e das limitações das soluções atuais em PLN. Os autores concluem que, apesar dos avanços, as promessas de maior explicabilidade e capacidade de raciocínio ainda não foram plenamente atingidas, reforçando a necessidade de aprofundar a integração entre paradigmas.

Por fim, o trabalho de Wang [Wang et al. 2024], propõe uma taxonomia abrangente que organiza as abordagens neuro-simbólicas em quatro eixos: integração neuro-simbólica, representação do conhecimento, incorporação de conhecimento e funcionalidade do sistema. A pesquisa oferece uma visão panorâmica da área, revisando diferentes métodos de integração entre dados e conhecimento, e apontando desafios e oportunidades para o futuro desenvolvimento de sistemas mais robustos e explicáveis.

Em conjunto, esses estudos reforçam a importância da utilização de modelos de classificação e taxonomias claras para a análise e o desenvolvimento de soluções em NeSy IA. Eles servem como referência para o presente trabalho, que busca justamente realizar a categorização de artigos da área a partir dos trabalhos de Bader & Hitzler e de Kautz.

3. Metodologia

Com o objetivo de trazer maior confiabilidade aos resultados e apresentá-los de forma sistematizada, nesta seção é apresentado o desenvolvimento da pesquisa, através da realização de busca, leitura e definição dos modelos de classificação dos artigos. A seção esta dividida em (i) mapeamento sistemático; (ii) classificação de Bader & Hitzler e (iii) classificação de Kautz.

3.1. Mapeamento

Para encontrar os artigos a serem estudados e classificados de acordo com as taxonomias anteriormente citadas, foi elaborado um mapeamento sistemático da literatura baseado no método de [Kitchenham and Charters 2007] para elencar os artigos da pesquisa. Foi utilizada a ferramenta Parsifal¹ para o auxílio ao mapeamento sistemático.

¹<https://parsif.al/> (Acessado em 17 de julho de 2025)

A realização do levantamento de artigos que utilizam técnica neuro-simbólicas ocorreu em partes (Figura 1) : (i) delimitação das strings de busca (SB); (ii) delimitação das bases de dados consultadas e (iii) critérios de exclusão (CE) de trabalhos.

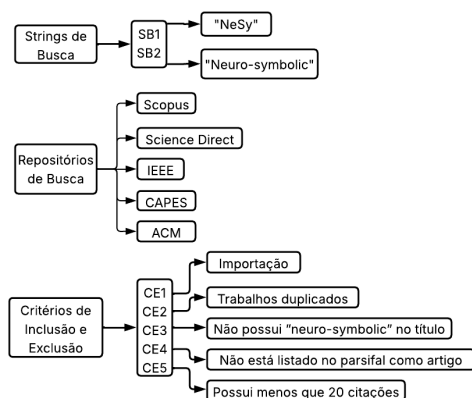


Figura 1. Etapas da execução do protocolo de mapeamento.

As strings de busca foram pensadas com o objetivo de pré selecionar trabalhos que utilizam NeSy AI, dessa forma, são apresentadas duas strings de busca: *Neuro-symbolic*, principal termo utilizado para identificar neuro-simbólicos em inteligência artificial e *NeSy*, termo abreviado de neuro-symbolic

A estratégia de busca foi definida visando identificar trabalhos publicados em repositórios de pesquisa. Os repositórios selecionados foram: *Scopus*, *Science Direct*, *IEEE*, *Coordenação de Aperfeiçoamento de Nível Superior (CAPES)* e *ACM*.

Os critérios de inclusão e exclusão permitiram selecionar os trabalhos que abordam a temática sobre NeSy AI, e foram definidos com base em:

- CE1 - Importação** Remove textos com erro ao importar para a ferramenta de auxílio;
- CE2 - Trabalhos Duplicados** Visa eliminar duplicidade de trabalhos;
- CE3 - Não possui "neuro-symbolic" no título** Remove os textos que não utilizam abordagem neuro-simbólica;
- CE4 - Não está listado no Parsifal como artigo** Elimina textos diferentes de artigos científicos, tais como capítulos de livros, notas e artigos de revisão;
- CE5 - Possui menos de 20 citações** Exclui da pesquisa, trabalhos com menos de vinte citações, considerando-as um indicativo de sua relevância acadêmica.

3.2. Classificação de Bader e Hitzler

[Bader and Hitzler 2005] propõem um modelo completo para classificar a integração entre a parte neural e simbólica dos sistemas. Os autores destacam questões ligadas ao uso do sistema e a linguagem simbólica adotada em 3 eixos, com 8 dimensões de análise cada, sendo elas:

Inter-relação: Descreve como as partes simbólicas e neurais de um sistema estão integradas;

1. Integrado vs Híbrido;

- Integrado: Sistema com um único componente neural principal, no qual o conhecimento simbólico é processado;

- Híbrido: Sistema caracterizado pela combinação de duas ou mais técnicas executadas em paralelo.
2. Neuronal vs Conexionista;
 - Neuronal: Preocupa-se em replicar comportamentos humanos, tendo inspiração biológica;
 - Conexionista: Foco em aplicações computacionais, sem necessariamente seguir a inspiração biológica.
 3. Local vs Distribuído;
 - Local: Representação explícita e mapeável, comum em sistemas baseados em proposições;
 - Distribuído: Representação implícita e difusa, típica de redes treinadas via backpropagation.
 4. Arquitetura Padrão vs Não Padrão;
 - Padrão: Favorece o aprendizado, apenas estruturas recursivas simples são utilizadas, como o backpropagation;
 - Não Padrão: Permite integrar representações simbólicas mais complexas, mas sacrifica as técnicas de aprendizado tradicionais.

Linguagem: refere-se à linguagem usada na parte simbólica do sistema;

1. Simbólica vs Lógica;
 - Simbólica: Representação explícita de conhecimento por meio de símbolos, pode-se citar strings, listas, árvores e grafos;
 - Lógica: Subconjunto do simbólico, refere-se especificamente ao uso de formalismos da lógica matemática.
2. Lógica Proposicional vs Primeira Ordem;
 - Lógica Proposicional: Manipula teorias com um número finito de variáveis. Exemplos incluem lógicas modais, temporais, não-monotônicas e outras não-clássicas;
 - Lógica de Primeira Ordem: Permite símbolos de função, pode envolver um número infinito de átomos.

Uso: Diz respeito a qual etapa do ciclo neuro-simbólico o sistema foca.

1. Representação: Forma como a informação estruturada é armazenada internamente por um sistema;
2. Extração: Recuperação de conhecimento simbólico que foi aprendido ou armazenado por um sistema;
3. Aprendizado: Capacidade de um sistema de ajustar seus parâmetros internos com base em dados, melhorando sua performance;
4. Raciocínio: Capacidade de um sistema de inferir novas informações a partir de conhecimento pré-existente utilizando mecanismos simbólicos.

3.3. Classificação de Kautz

O modelo de classificação de Kautz, proposto em [Kautz 2021], mostrou uma estagnação no avanço da IA dado a necessidade da implementação conjunta de raciocínio simbólico e neural/conexionista. O trabalho elenca seis designs de arquitetura para sistemas neuro-simbólicos, sendo eles:

Symbolic→Neuro→Symbolic Entradas simbólicas, processamento por redes neurais, e resultados simbólicos;

Symbolic[Neuro] Um resolvidor simbólico externo usa componentes neurais como funções auxiliares;

Neuro | Symbolic Possui duas etapas principais;

1. A rede neural transforma dados brutos em estruturas simbólicas;
2. Um sistema simbólico faz o raciocínio.

Neuro: Symbolic → **Neuro** Uma arquitetura neural, mas treinada com base em pares de entrada/saída gerados por um sistema simbólico;

Neuro-{Symbolic} Estruturas simbólicas são traduzidas para pesos, conexões ou templates da rede neural;

Neuro[Symbolic] Uma fusão profunda dos dois paradigmas. Segundo Kautz, é o sistema neuro-simbólico com o maior potencial de combinar as forças neural e simbólica.

4. Resultados

Para facilitar a compreensão dos dados coletados e análises de forma mais, realizou-se uma análise detalhada da distribuição dos trabalhos encontrados, destacando os principais repositórios utilizados e a eficácia das estratégias de busca aplicadas.

4.1. Mapeamento

Os resultados foram satisfatórios (Figura 2), onde tem-se os repositórios e a quantidade de trabalhos retornados com o uso de cada string de busca.

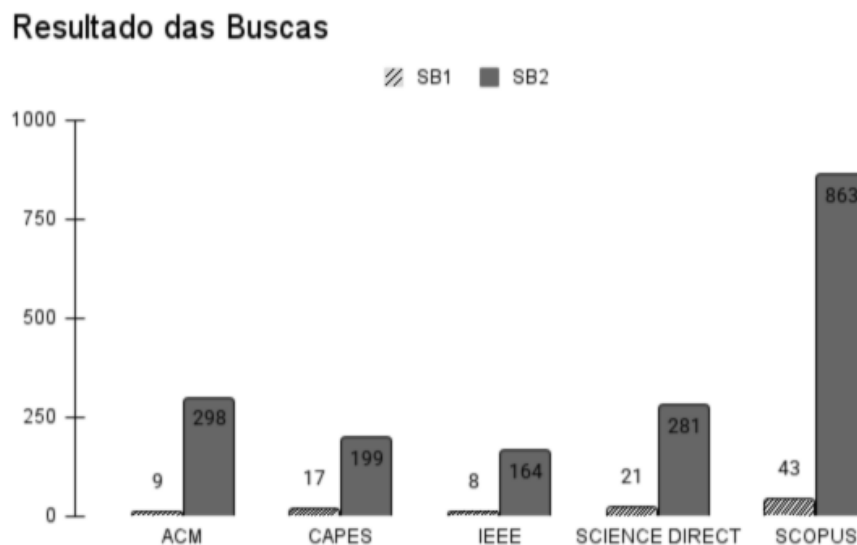


Figura 2. Resultados das buscas por repositório.

Um ponto de destaque notado na busca, trata-se do volume expressivo de resultados provenientes do repositório SCOPUS, indicando este como um repositório estratégico para buscas e pesquisas na área de NeSy.

A soma de todos os resultados obtidos com as buscas nos repositórios totalizou 1903 textos. Com a aplicação dos dois primeiros critérios, houve uma redução de 1903 para 1396. O terceiro critério eliminou mais de 1000 manuscritos, de forma a restarem 384, dos quais 10 foram selecionados após a execução do quarto e quinto critérios (Tabela1).

ID	Título	Referência
1	Neuro-Symbolic Models for Sentiment Analysis	[Kocoń et al. 2022]
2	autoBOT: evolving neuro-symbolic representations for explainable low resource text classification	[Škrlj B. et al. 2021]
3	PIGLeT: Language Grounding Through Neuro-Symbolic Interaction in a 3D World	[Zellers et al. 2022]
4	Dynamic Neuro-Symbolic Knowledge Graph Construction for Zero-Shot Commonsense Question Answering	[Bosselut et al. 2021]
5	Improving Coherence and Consistency in Neural Sequence Models with Dual-System, Neuro-Symbolic Reasoning	[Nye et al. 2021]
6	Neuro-Symbolic Visual Reasoning: Disentangling “Visual” from “Reasoning”	[Amizadeh et al. 2020]
7	Semantic Probabilistic Layers for Neuro-Symbolic Learning	[Ahmed et al. 2022]
8	Learning Neuro-Symbolic Relational Transition Models for Bilevel Planning	[Chitnis et al. 2022]
9	Neuro-Symbolic Hierarchical Rule Induction	[Glanois et al. 2022]
10	NS3D: Neuro-Symbolic Grounding of 3D Objects and Relations	[Hsu et al. 2023]

Tabela 1. Artigos selecionados para classificação.

4.2. Classificações

A seguir as classificações são realizadas, levando em consideração as taxonomias de Bader & Hitler e de Kautz. Vale ressaltar o artigo (ID: 1) escrito por [Kocoń et al. 2022], apresenta mais de um modelo que utiliza neuro-simbolismo, neste caso, o escolhido para classificação foi o “Tailored KEPLER”, uma vez que este é destacado nas conclusões do trabalho original.

Com a análise finalizada (Quadro1), é possível perceber alguns fatos, ao que diz respeito às oito dimensões, distribuídas entre os eixos de inter-relação, linguagem e uso. Apresenta-se neste estudo, um maior número de abordagens híbridas, com a totalidade do uso de sistemas conexionistas distribuídos, com arquitetura não padrão, a linguagem simbólica é unânime, usada majoritariamente com a proposicional. O uso dos sistemas, em sua maioria é completo, realizando representação, extração, aprendizado e raciocínio.

Em relação ao modelo proposto por Kautz, teve-se a maior ocorrência de sistemas *Neuro | Symbolic* com sete textos, *Neuro_{Symbolic}* totaliza dois, e *Neuro:Symbolic* → *Neuro* com somente uma implementação. É interessante visualizar que, *Neuro[Symbolic]* continua sem apresentar implementações.

Autor	Classificação	1	2	3	4	5	6	7	8	9	10
Bader & Hitzler, 2005	Híbrido		✓	✓	✓	✓		✓		✓	✓
	Integrado	✓					✓		✓		
	Neural										
	Conexionista	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Local										
	Distribuída	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Arq. Padrão										
	Arq. Não Padrão	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Simbólica	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Lógica										
	Proposicional	✓	✓		✓	✓	✓		✓	✓	✓
	Primeira-ordem			✓				✓			
	Representação	✓	✓	✓	✓		✓	✓	✓	✓	✓
	Extração	✓	✓	✓	✓	✓		✓	✓	✓	✓
Kautz, 2021	Aprendizado	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Raciocínio	✓	✓	✓	✓	✓		✓	✓	✓	✓
	Symbolic→Neuro→symbolic										
	Symbolic[Neuro]										
	Neuro Symbolic		✓	✓	✓	✓		✓		✓	✓
	Neuro:Symbolic→Neuro						✓				
	Neuro_{Symbolic}	✓							✓		
	Neuro[Symbolic]										

Quadro 1: Classificação dos artigos segundo os critérios de Bader & Hitzler (2005) e Kautz (2021).

5. Conclusão

Este artigo apresentou a busca e classificação de trabalhos recentes na área de NeSy AI, utilizando duas taxonomias consolidadas. O estudo identificou padrões nas implementações, destacando que a maioria dos sistemas combina duas ou mais técnicas em paralelo, utilizando redes neurais conexionistas com representações implícitas e distribuídas, que integram símbolos como strings e outras estruturas, manipulando um número finito de variáveis. Observou-se que os sistemas geralmente recuperam conhecimento simbólico, ajustam seus parâmetros para melhorar desempenho, e realizam raciocínio simbólico para inferir e armazenar novas informações. Também foi confirmada a ausência de implementações do design *Neuro[Symbolic]* proposto por Kautz, considerado o mais promissor para a integração profunda dos paradigmas neural e simbólico.

O trabalho ressalta a importância da utilização de métodos de comparação entre abordagens neuro-simbólicas, que possibilitam uma análise prévia e orientam a escolha de estratégias focadas em objetivos específicos, como representação ou raciocínio. Como direcionamento para futuros estudos, discute-se a criação de um sistema de classificação autoral, que não apenas categorize diferentes propostas neuro-simbólicas e formas de integração, mas que também leve em conta as técnicas utilizadas, as linguagens de implementação e as áreas de aplicação desses sistemas.

Agradecimentos: Este trabalho tem financiamento do projeto: 444568/2024-7, CNPq/MCTI/FNDCT Nº 22/2024, *Programa Conhecimento Brasil – Apoio a Projetos em Rede com Pesquisadores Brasileiros no Exterior*.

Referências

- Ahmed, K., Teso, S., Chang, K.-W., den Broeck, G. V., and Vergari, A. (2022). Semantic probabilistic layers for neuro-symbolic learning. *Advances in Neural Information Processing Systems*.
- Amizadeh, S., Palangi, H., Polozov, O., Huang, Y., and Koishida, K. (2020). Neuro-symbolic visual reasoning: Disentangling "visual" from "reasoning". *37th International Conference on Machine Learning, ICML 2020*.
- Bader, S. and Hitzler, P. (2005). Dimensions of neural-symbolic integration - a structured survey. *arXiv:cs/0511042*.
- Bosselut, A., Bras, R. L., and Choi, Y. (2021). Dynamic neuro-symbolic knowledge graph construction for zero-shot commonsense question answering. *35th AAAI Conference on Artificial Intelligence, AAAI 2021*.
- Chitnis, R., Silver, T., Tenenbaum, J. B., Lozano-Perez, T., and Kaelbling, L. P. (2022). Learning neuro-symbolic relational transition models for bilevel planning. *IEEE International Conference on Intelligent Robots and Systems*.
- Dennis, L., Farrell, M., and Fisher, M. (2023). Developing multi-agent systems with degrees of neuro-symbolic integration. *arXiv:2305.11534v1*.
- Glanois, C., Feng, X., Jiang, Z., Weng, P., Zimmer, M., Li, D., and Liu, W. (2022). Neuro-symbolic hierarchical rule induction. *Proceedings of Machine Learning Research*.
- Hagos, Haileselassie, D., Rawat, and B., D. (2024). Neuro-symbolic ai for military applications. *IEEE Transactions on Artificial Intelligence*.
- Hamilton, K., Nayak, A., Božić, B., and Longo, L. (2024). Is neuro-symbolic ai meeting its promises in natural language processing? a structured review. *Semantic Web Volume 15, Issue 4: Neuro-Symbolic Artificial Intelligence and the Semantic Web*.
- Hsu, J., Mao, J., and Wu, J. (2023). Ns3d: Neuro-symbolic grounding of 3d objects and relations. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- Kautz, H. A. (2021). The third ai summer: Aaai robert s. engelmore memorial lecture. *AI Magazine 43: 105–125*.
- Kitchenham, B. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. *Technical Report EBSE 2007-001*.
- Kocoń, J., Baran, J., Gruza, M., Janz, A., Kajstura, M., Kazienko, P., Korczyński, W., Miłkowski, P., Piasecki, M., and Szolomicka, J. (2022). Neuro-symbolic models for sentiment analysis. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.
- Nye, M., Tessler, M. H., Tenenbaum, J. B., and Lake, B. M. (2021). Improving coherence and consistency in neural sequence models with dual-system, neuro-symbolic reasoning. *Advances in Neural Information Processing Systems*.

- Saker, M. K., Zhou, L., Eberhart, A., and Hitzler, P. (2021). Neuro-symbolic artificial intelligence current trends. *arXiv:2105.05330v1*.
- van Bekkum, M., de Boer, M., van Harmelen, F., Meyer-Vitali, A., and ten Teije, A. (2021). Modular design patterns for hybrid learning and reasoning systems: a taxonomy, patterns and use cases.
- Wang, W., Yang, Y., and Wu, F. (2024). Towards data-and knowledge-driven artificial intelligence: A survey on neuro-symbolic computing.
- Zellers, R., Holtzman, A., Peters, M., Mottaghi, R., Kembhavi, A., Farhadi, A., and Choi, Y. (2022). Piglet: Language grounding through neuro-symbolic interaction in a 3d world. *ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference*.
- Škrlj B., M., M., N., L., and S., P. (2021). autobot: evolving neuro-symbolic representations for explainable low resource text classification. *Springer Machine Learning Volume 110, Issue 5, May 2021, Pages 989-1028*.