Sobre a Interação entre Inteligência Artificial Explicável (XAI) e Ontologias: Uma Revisão Quasi-Sistemática da Literatura

Lucas G. Maddalena[1,*], Fernanda Baião[1], Tiago Prince Sales[2] and Giancarlo Guizzardi[2]

[1]*Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rua Marquês de São Vicente, 225, Rio de Janeiro, 22451-900, Brasil*

[2]*Semantics, Cybersecurity & Services, University of Twente, Enschede, The Netherlands*

Abstract

Explainable Artificial Intelligence (XAI) seeks to reconcile the predictive power of modern Artificial Intelligence (AI) models—typically based purely on data-driven learning—with the human need for transparent and trustworthy decisions. This paper presents a quasi-systematic literature review, grounded in the PRISMA protocol principles [1], which examines the interplay between ontologies—formal, machine-readable representations of domain knowledge—and contemporary XAI techniques. A total of 440 records were retrieved from the Scopus database (2018–2025), which were screened and filtered according to predefined inclusion criteria. The final corpus was complemented by three seminal works recognized for the depth of their conceptual contributions, totaling 31 articles. The analysis aims to address three central questions: (1) how the notion of "explanation" is defined and operationalized in the XAI literature; (2) in which ways ontologies are employed as semantic artifacts in XAI systems—from term mapping to counterfactual reasoning; and (3) what impact ontological grounding has on the expressiveness of explanations and on user understanding. The results indicate that the concept of explanation is addressed in a heterogeneous manner—encompassing mechanistic, contrastive, and social perspectives—while ontologies are often used for domain-specific visualization and contextualization, with foundational ontologies still being underexplored. Ontological anchoring of explanations tends to enrich their semantic depth and shows potential to enhance user understanding, although empirical validation remains limited. The paper concludes with a discussion of emerging neuro-symbolic approaches and suggests future directions for integrating well-founded ontologies into explainable AI frameworks.

Keywords

Explainable Artificial Intelligence, Ontologies, XAI, Foundational Ontologies, Explainability, Literature Review

1. Introdução

Nos últimos anos, o campo da Inteligência Artificial (IA) tem vivenciado um notável aumento de popularidade. Embora existam várias definições sobre o que realmente constitui IA, adota-se neste trabalho a definição de Sheikh et al., que a descreve como sistemas que exibem comportamento inteligente ao analisar seu ambiente e tomar ações — com algum grau de autonomia — para alcançar objetivos específicos [2].

Contudo, as abordagens modernas de IA — especialmente aquelas baseadas em Aprendizado de Máquina (*Machine Learning*) e, em particular, em Aprendizado Profundo (*Deep Learning*) — são essencialmente orientadas por dados, baseando-se quase exclusivamente em padrões extraídos de grandes volumes de dados, que são considerados uma inteligência de conhecimento sub-simbólico, em detrimento da utilização de conhecimento externo codificado de forma explícita. Neste contexto, é notório que tais padrões apreendidos podem ser enviesados, não representativos e desprovidos de ancoragem no mundo real. Como resultado, a última década testemunhou tanto a ascensão das Redes Neurais

Proceedings of the 18th Seminar on Ontology Research in Brazil (ONTOBRAS 2025) and 9th Doctoral and Masters Consortium on Ontologies (WTDO 2025), São José dos Campos (SP), Brazil, September 29 – October 02, 2025.

*Autor correspondente.

lucasmadda@aluno.puc-rio.br (L. G. Maddalena); fbaiao@puc-rio.br (F. Baião); t.princesales@utwente.nl (T. P. Sales); g.guizzardi@utwente.nl (G. Guizzardi)

0000-0001-6411-4259 (L. G. Maddalena); 0000-0001-7932-7134 (F. Baião); 0000-0002-5385-5761 (T. P. Sales); 0000-0002-3452-553X (G. Guizzardi)

Profundas — arquiteturas com centenas de camadas e milhões de parâmetros — quanto uma conscientização crescente de que tais arquiteturas resultam em modelos opacos, caracterizando-os como sistemas do tipo "caixa-preta" [3]. Diante desse cenário, a Inteligência Artificial Explicável (XAI) tem se consolidado como uma demanda central na academia e na indústria. À medida que sistemas de IA são aplicados em domínios críticos — como Saúde, Finanças, Sistemas Autônomos e Segurança —, torna-se essencial garantir que as partes interessadas compreendam o raciocínio subjacente dos modelos, construam confiança nos resultados produzidos e avaliem a robustez das previsões frente a entradas fora da distribuição, isto é, situações em que o modelo recebe dados substancialmente diferentes daqueles usados em seu treinamento. Exemplos típicos incluem exames clínicos realizados com equipamentos distintos ou em populações de pacientes com características demográficas ausentes no conjunto original de dados. Nesses casos, mesmo modelos altamente precisos podem apresentar degradação de desempenho ou produzir saídas enganosas, evidenciando a necessidade de explicações que revelem suas limitações. No entanto, embora essa necessidade venha sendo amplamente reconhecida, ainda há considerável divergência sobre o que constitui uma explicação válida. Em muitos casos, os métodos de XAI revelam apenas correlações estatísticas aprendidas a partir dos dados de treinamento, sem oferecer verdadeiro entendimento sobre o funcionamento do modelo, podendo inclusive reforçar vieses cognitivos ao fornecer justificativas superficiais.

O conhecimento externo explicitamente estruturado pode ser representado por ontologias — formalmente definidas como "uma especificação formal e explícita de uma conceitualização compartilhada" [4] — as quais oferecem uma fonte complementar de informação aos métodos puramente baseados em dados. Enquanto grande parte das técnicas de XAI se limita à visualização de padrões extraídos dos dados, as ontologias permitem fundamentar o comportamento dos modelos em conhecimento estruturado e legível por máquina, com potencial para mitigar vieses, ampliar a robustez preditiva e produzir explicações mais informativas e semanticamente válidas. Ao explicitar relações, categorias ontológicas e restrições do domínio, essas estruturas favorecem a geração de explicações alinhadas com o mundo real, promovendo maior transparência e apoio à decisão. Nesse sentido, a integração de modelos conceituais e aprendizado de máquina — como discutido em [5] — oferece uma via promissora para alinhar abstrações semânticas com ciclos de desenvolvimento em ciência de dados. Essa sinergia entre modelagem ontológica e aprendizado de máquina contribui não apenas para aumentar a expressividade das explicações, mas também para ancorá-las em representações cognitivamente adequadas e relevantes para os usuários finais.

Embora já existam revisões relevantes sobre o tema, como a de Ali et al. [6], que apresenta um mapeamento abrangente das técnicas e métricas de XAI, e a de Rajabi and Etminani [7], que sistematiza o uso de grafos de conhecimento em diferentes estágios de modelos explicáveis, ambas seguem direções distintas da proposta aqui desenvolvida. A primeira centra-se em categorizar métodos e ferramentas disponíveis para XAI em geral, sem adentrar especificamente no papel de ontologias; a segunda foca principalmente em grafos de conhecimento, mencionando ontologias apenas de forma tangencial, sem discutir sua distinção conceitual nem seu potencial quando fundamentadas em ontologias de topo. Em contraste, o presente trabalho busca examinar como ontologias — enquanto artefatos semânticos formais e bem fundamentados — têm sido integradas a sistemas de XAI, enfatizando seus efeitos na qualidade e expressividade das explicações. Dessa forma, nossa revisão avança em uma direção ainda pouco explorada, complementando os panoramas já existentes e preenchendo uma lacuna sobre o uso de ontologias propriamente ditas na explicabilidade de modelos de IA.

Com base nessa perspectiva, este trabalho realiza uma revisão quasi-sistemática da literatura com o objetivo de examinar como ontologias e outras formas de representação do conhecimento vêm sendo incorporadas a técnicas contemporâneas de XAI, com ênfase especial em abordagens neurossimbólicas que integram raciocínio simbólico e aprendizado sub-simbólico, e endereçando três questões centrais de pesquisa: (QP1) Como o termo "explicação" é definido e operacionalizado na literatura atual sobre XAI?; (QP2) De que formas as ontologias são utilizadas como artefatos semânticos em sistemas de XAI?; e (QP3) Como a fundamentação de explicações em ontologias afeta sua expressividade e a compreensão por parte dos usuários?

Inicialmente, o artigo delimita os conceitos fundamentais que sustentam a pesquisa, com destaque para

a Inteligência Artificial Explicável, a função das ontologias como artefatos semânticos e os paradigmas emergentes de IA neurossimbólica. Essa fundamentação teórica estabelece o alicerce para a metodologia apresentada na Seção 3, na qual são descritos o desenho da pesquisa, os critérios de inclusão e exclusão, a estratégia de busca utilizada e os procedimentos de triagem e seleção bibliográfica. A Seção 4 sintetiza os principais achados da revisão, identificando padrões recorrentes e lacunas existentes na literatura sobre a integração entre ontologias e XAI. São discutidos criticamente os tipos de ontologias utilizados, os mecanismos de explicação empregados, bem como os efeitos observados na transparência e na compreensão por parte dos usuários. Por fim, a Seção 5 apresenta considerações finais sobre as implicações dos resultados obtidos e propõe direções promissoras para pesquisas futuras.

2. Referencial Teórico

A Inteligência Artificial Explicável (XAI) surgiu como resposta direta à crescente opacidade dos modelos modernos de aprendizado de máquina — particularmente das redes neurais profundas, cujas arquiteturas podem abranger centenas de camadas e milhões de parâmetros — tornando-os sistemas "caixa-preta" de difícil interpretação [3]. Paradigmas iniciais de IA, como os sistemas especialistas baseados em regras, forneciam caminhos decisórios inerentemente transparentes, mas as abordagens orientadas por dados abrem mão dessa clareza em prol do poder preditivo. Como observam Confalonieri et al. [8], essa mudança deu origem a preocupações cognitivas e sociais urgentes: os envolvidos precisam não apenas saber o que o modelo prediz, mas também entender o porquê — de modo que as explicações estejam alinhadas com noções humanas de causalidade e responsabilidade. Mais recentemente, Ali et al. [6] destacam que a XAI é fundamental para a construção de uma IA confiável (Trustworthy AI) — garantindo transparência, confiabilidade e justiça em domínios críticos como saúde e finanças. No entanto, como é alertado por Mittelstadt et al. [9], muitas das chamadas "explicações" são pouco mais que modelos substitutos simplificados, reaproveitados da modelagem científica, mas sem estarem fundamentados nas práticas contrastivas, seletivas e situadas socialmente que caracterizam as explicações genuinamente humanas.

Para trazer ordem a esse campo complexo, taxonomias de métodos XAI geralmente distinguem entre modelos ditos "intrinsecamente transparentes" e aqueles que exigem interpretação *post-hoc*. A primeira categoria — que inclui regressões (lineares, logísticas, entre outras), árvores de decisão e até mesmo ferramentas baseadas em modelos substitutos como LIME [10] e SHAP [11] — oferece rastreabilidade direta por meio de coeficientes ou regras explícitas, mas ainda assim permanece totalmente orientada por dados e vulnerável a extrapolações enganosas em entradas fora da distribuição [3]. Por outro lado, as técnicas *post-hoc* se dividem em abordagens independentes de modelo (modelo como caixa-preta) e métodos específicos de modelo que aproveitam detalhes arquiteturais (ex: mapas de saliência, propagação de relevância por camadas). As explicações também podem ser classificadas como locais — que elucidam previsões individuais — ou globais — que resumem o comportamento geral do modelo. Embora essas distinções auxiliem na escolha e avaliação dos métodos, Mittelstadt et al. [9] nos lembram que um foco estreito em aproximações funcionais corre o risco de negligenciar as dimensões normativas mais profundas da explicação — como a necessidade de contrafactuais contrastivos, relevância contextual e interação dialógica. Abordar essas lacunas é central para avançar a XAI de um conjunto de ferramentas exclusivamente técnicas para uma disciplina centrada no ser humano, que de fato capacite tanto usuários quanto reguladores.

Com base nessas fundações, Schwalbe and Finzel [12] propõem uma taxonomia abrangente de métodos XAI, ilustrada na 1, que organiza sistematicamente as características explicativas em dimensões inter-relacionadas. Partindo dos requisitos de entrada (incluindo acesso ao modelo, dados relevantes, feedback do usuário e informações contextuais), a taxonomia distingue os métodos por portabilidade (agnóstico vs. específico de modelo) e localidade (explicações globais versus locais). Em seguida, define critérios de interatividade (fluxos de trabalho exploratórios ou corretivos), restrições matemáticas (linearidade, monotonicidade, satisfatibilidade, limites de iteração e tamanho do modelo) e tipos de saída (baseados em exemplos, contrastivos/contrafactuais, guiados por protótipos, atribuições de características, baseados

em regras, redução de dimensionalidade, gráficos/diagramas). Por fim, aborda fatores de apresentação — modalidade (visual, verbal, auditiva), nível de abstração, acessibilidade, granularidade da informação e consciência de privacidade — oferecendo, assim, um arcabouço estruturado para seleção e avaliação de técnicas XAI conforme os requisitos específicos de uso.

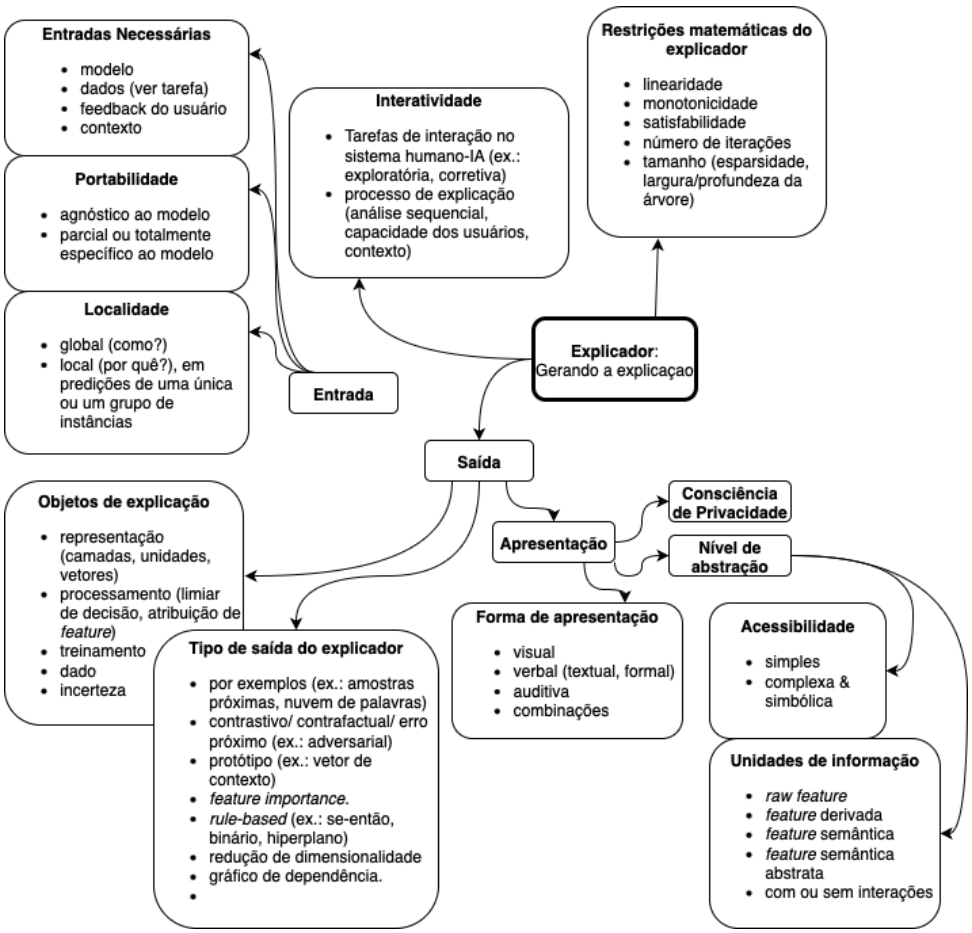

Figura 1: Aspectos taxonômicos de um explanador, conforme proposto por [12]

A partir dessa taxonomia e do reconhecimento das limitações das explicações puramente orientadas por dados, a IA Neurossimbólica surge como uma promissora terceira onda da Inteligência Artificial, unindo a capacidade de reconhecimento de padrões do aprendizado de máquina (profundo) ao rigor e interpretabilidade do raciocínio simbólico [13, 14]. Enquanto redes sub-simbólicas se destacam na extração de características latentes a partir de grandes conjuntos de dados, mas oferecem pouco em termos de justificativas compreensíveis para humanos, os sistemas neurossimbólicos incorporam representações formais de conhecimento — ontologias, regras lógicas ou restrições probabilísticas — diretamente nas arquiteturas neurais ou em paralelo a elas. Essa fusão fundamentada permite que os modelos realizem aprendizado por gradiente em tarefas perceptivas ao mesmo tempo em que sustentam inferência simbólica estruturada, raciocínio contrafactual contrastivo e generalização composicional. Ao unir esses paradigmas, a IA Neurossimbólica pretende combinar alto desempenho preditivo com caminhos decisórios intrinsecamente explicáveis, lançando as bases para abordagens XAI que realmente abordem as dimensões normativas, seletivas e interativas da explicação humana.

Essa convergência entre o rigor simbólico e o aprendizado sub-simbólico ressalta a necessidade de ontologias de fundamentação de alta qualidade como alicerce dos sistemas explicáveis. Estruturas como a Unified Foundational Ontology (UFO)[15], a DOLCE [16] e a Basic Formal Ontology (BFO) [17] oferecem conjuntos gerais de categorias e relações, apoiados em compromissos filosóficos claros e metodologias consolidadas. Ao ancorar o conhecimento de domínio nessas ontologias e ao empregar

linguagens de modelagem conceitual como a OntoUML, garante-se que o componente simbólico seja simultaneamente semanticamente rico e formalmente consistente. Essas ontologias impõem restrições à inferência neural e possibilitam explicações rastreáveis a conceitos compartilhados e legíveis por humanos, reduzindo a lacuna de interpretabilidade inerente às abordagens exclusivamente orientadas por dados.

3. Metodologia

Este estudo adota uma abordagem de *revisão quasi-sistemática* da literatura, guiada por princípios do protocolo PRISMA [1], com o objetivo de assegurar transparência e reprodutibilidade na seleção dos trabalhos analisados. Foram empregadas boas práticas metodológicas amplamente reconhecidas, tais como: definição antecipada de critérios de inclusão e exclusão; construção de uma estratégia de busca estruturada; e registro visual do processo de triagem por meio de fluxograma. A revisão foi orientada por três perguntas centrais de pesquisa:

- **QP1**: Como o termo "explicação" é definido e operacionalizado na literatura atual sobre XAI?
- **QP2**: De que formas as ontologias são utilizadas como artefatos semânticos em sistemas de XAI?
- **QP3**: Como a fundamentação de explicações em ontologias afeta sua expressividade e a compreensão por parte dos usuários?

Com relação à QP1, buscou-se compreender, especificamente, quais dimensões de explicação são enfatizadas pelos autores e quão consistentes essas definições são entre domínios e venues. Essas dimensões incluem perspectivas mecanicistas (como entradas específicas afetam as saídas de um modelo), contrastivas (explicações do tipo "por que A e não B?") e sociais (que consideram o perfil, contexto e necessidades do usuário final) [12]. Para a QP2, investigou-se quais tipos de ontologias (de fundamentação vs. específicas de domínio) aparecem com maior frequência e quais papéis desempenham (ex.: fundamentação de termos, estruturação causal, raciocínio contrafactual). Finalmente, na QP3, buscou-se compreender se explicações baseadas em ontologias demonstram maior riqueza semântica (relações explícitas, restrições) e se há evidências empíricas de que os usuários as consideram mais transparentes ou acionáveis.

A base de dados Scopus foi consultada utilizando a seguinte string de busca: `TITLE-ABS-KEY((“explainable ai” OR xai OR explainability OR interpretability OR “neuro-symbolic”) AND (ontology OR ontologies OR ontological OR “foundational ontology” OR ufo OR ontouml))`.

Com o intuito de concentrar a análise em contribuições mais consolidadas no campo da XAI, foram considerados apenas artigos publicados em sua versão final a partir de 2018 — marco temporal a partir do qual o termo "XAI" passou a figurar com maior destaque em títulos, resumos e palavras-chave.

A consulta inicial retornou 440 registros. Primeiramente, removemos 386 itens que não eram artigos completos (como cartas, atas de conferências, editoriais), restando 54 artigos para triagem por título, resumo e palavras-chave. Nessa fase, excluímos 17 estudos sem ligação substancial com IA explicável, ontologias ou sistemas de representação do conhecimento, além de outros 9 artigos que mencionavam XAI apenas de forma tangencial, sem discussão explícita sobre mecanismos de explicação, fundamentos ontológicos ou integração sistemática de artefatos semânticos. Isso resultou em 28 artigos selecionados após a triagem.

Como mostrado na Figura 2, recuperamos 440 registros da Scopus, dos quais 386 foram removidos por não serem pertinentes. Os 54 restantes foram avaliados com base em título, resumo e, quando necessário, no texto completo; nessa etapa, 17 foram excluídos por ausência de vínculo com XAI ou ontologias, e 9 por não abordarem mecanismos de explicação ou fundamentos ontológicos. Reconhecendo, entretanto, que nossas questões de pesquisa demandavam uma análise aprofundada das estruturas ontológicas aplicadas em XAI, o corpus foi complementado com três trabalhos seminais previamente conhecidos por suas definições rigorosas e tratamento detalhado do papel das ontologias na explicação [18, 19, 20].

Assim, a amostra final totalizou 31 artigos, submetidos à extração e síntese detalhadas para responder às três perguntas de pesquisa propostas.

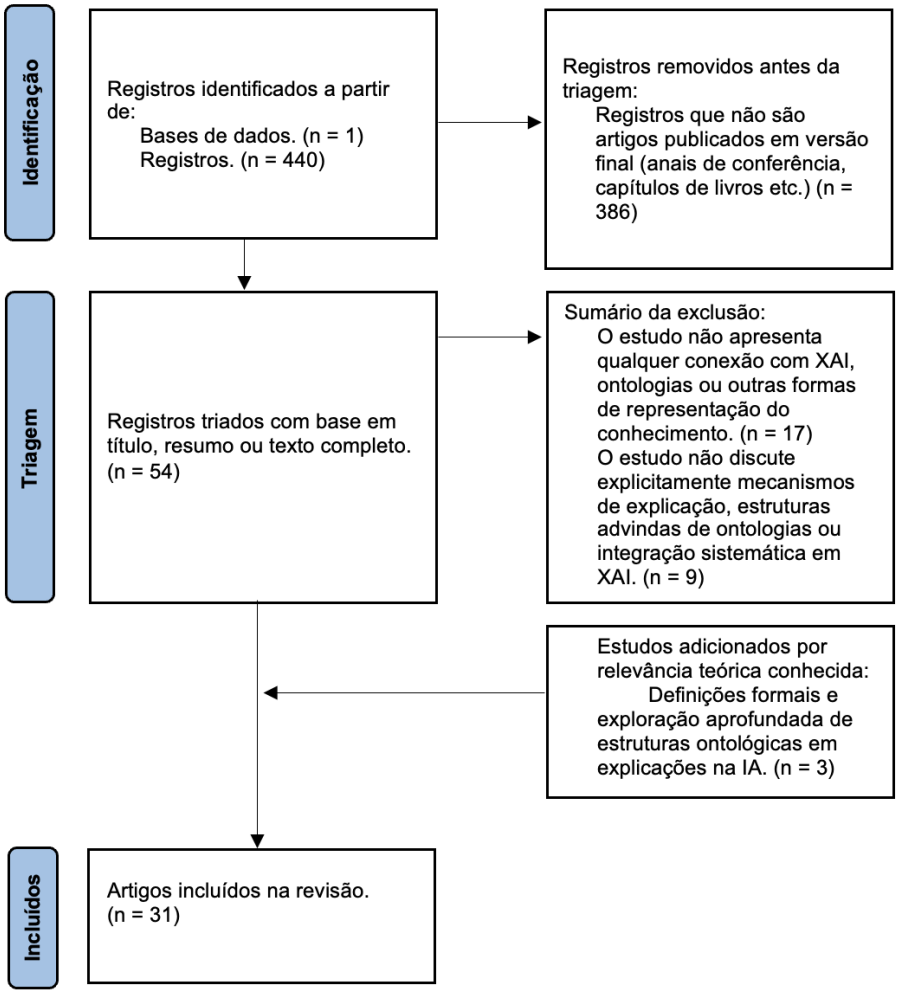

Figura 2: Fluxograma do processo de seleção dos estudos. Fonte: Elaborado pelos autores.

4. Resultados

A Tabela 1 resume as publicações selecionadas e suas respectivas contribuições para cada uma das três perguntas de pesquisa (QP1: definição e operacionalização de "explicação" no contexto de XAI; QP2: papel das ontologias como artefatos semânticos; e QP3: impacto da fundamentação ontológica na expressividade das explicações e na compreensão por parte dos usuários), indicando se a publicação apresentou contribuição relevante ("Sim") ou não ("Não") e assinalando com "*" os casos em que observações ou nuances adicionais são discutidas posteriormente. Conforme evidenciado, alguns trabalhos, como [18, 19, 20], oferecem contribuições significativas para todas as perguntas de pesquisa, ao passo que estudos como [21] e [22] não abordam diretamente nenhuma das questões delineadas.

4.1. Análise da QP1: Definição e operacionalização de "explicação"em XAI

A maioria das publicações selecionadas aborda, de forma direta ou indireta, a primeira questão de pesquisa (QP1), relacionada à definição e operacionalização do conceito de explicação em Inteligência Artificial Explicável (XAI). Entretanto, alguns estudos apresentam limitações quanto à profundidade dessa discussão. Por exemplo, Sun et al. [21] apenas menciona a falta de interpretabilidade em modelos

Tabela 1
Artigos selecionados e contribuições para as Questões de Pesquisa endereçadas no presente estudo.

Citação	Referência	Ano	QP1	QP2	QP3
[18]	Guizzardi and Guarino	2024	Sim	Sim	Sim
[19]	Confalonieri and Guizzardi	2023	Sim	Sim	Sim
[20]	Chari et al.	2020	Sim	Sim	Sim
[21]	Sun et al.	2022	Não	Não	Não
[22]	Lu et al.	2023	Sim	Não	Não
[23]	Li et al.	2025	Sim	Sim*	Não
[24]	Liu et al.	2022	Sim	Sim	Sim
[25]	Perera and Liu	2024	Sim	Sim*	Não*
[26]	Ma et al.	2025	Sim*	Sim*	Não
[27]	Wei et al.	2024	Sim*	Não	Não
[28]	Zhao et al.	2024	Sim	Sim	Sim
[29]	Rasbach et al.	2024	Sim	Não*	Não
[30]	Dubey et al.	2025	Sim	Não	Não
[31]	Zhao et al.	2022	Sim	Sim	Não
[32]	Zhan et al.	2025	Sim	Sim	Sim
[33]	Sung et al.	2024	Sim	Sim	Sim
[34]	Li et al.	2025	Sim	Sim	Sim
[35]	Zheng et al.	2024	Sim	Sim	Não
[36]	Chen et al.	2018	Não*	Sim*	Sim*
[37]	Canito et al.	2021	Sim*	Sim*	Sim*
[38]	Meghraoui et al.	2025	Não*	Sim	Sim
[39]	Guan et al.	2024	Sim	Não	Não
[40]	Basakin et al.	2024	Sim	Sim	Sim
[41]	Li et al.	2023	Sim	Não	Sim
[42]	Kaptein et al.	2022	Sim	Sim	Não
[43]	Smedley et al.	2021	Sim	Sim	Sim
[44]	Ruiz-Arenas et al.	2024	Sim	Sim	Não
[45]	Seninge et al.	2021	Sim	Sim	Sim
[46]	Singha et al.	2023	Sim	Não	Não
[47]	Askland et al.	2021	Sim*	Sim	Sim
[48]	Townsend et al.	2024	Sim	Não	Sim
	Total de "Sim"		**27**	**23**	**17**

de IA do tipo caixa-preta, sem aprofundar nos tipos de explicação ou modelos adotados na literatura. De forma semelhante, Ma et al. [26] reconhece a importância da interpretabilidade, mas não a relaciona com as dimensões conceituais da explicação tratadas no campo de XAI. O estudo de Wei et al. [27], embora utilize recursos funcionais para predizer variantes genéticas patogênicas, também não define explicitamente o que se entende por "explicação" nem discute suas possíveis dimensões.

Já os trabalhos de Chen et al. [36] e Meghraoui et al. [38] exemplificam abordagens que, embora não se identifiquem diretamente com o campo de XAI como hoje é concebido, contribuem para a interpretabilidade dos modelos por meio da incorporação de conhecimento estruturado. No primeiro caso, são integrados recursos como WordNet [49] e UMLS (Unified Medical Language System) [50] às análises de embeddings. No segundo, adota-se uma abordagem neurossimbólica aplicada à previsão de produtividade agrícola, cuja natureza simbólica intrínseca tende a favorecer a explicabilidade do sistema, ainda que sem uma formalização explícita do conceito de explicação.

No estudo de Canito et al. [37], é apresentada uma revisão sistemática sobre evolução ontológica com restrição temporal no domínio de manutenção preditiva. Embora não esteja diretamente inserido no escopo da XAI, o trabalho oferece uma contribuição relevante à operacionalização de explicações em contextos que envolvem conhecimento estruturado e dinâmico ao longo do tempo. O modelo proposto para representar a evolução de ontologias com validade temporal relaciona-se ao desafio de justificar

e rastrear mudanças no comportamento de sistemas inteligentes, o que é essencial para explicações retrospectivas e prospectivas. A marcação (*) na Tabela 1, nesse caso, reflete o fato de tratar-se de uma revisão sistemática, cuja abrangência metodológica contribui indiretamente para a semântica da explicação por meio de mecanismos ontológicos, no caso em questão sensíveis especificamente ao aspecto temporal.

Já no trabalho de Askland et al. [47], embora não haja definição ou operacionalização explícita de "explicação" nas dimensões mais comuns da literatura — como as perspectivas mecanicista, contrastiva ou social —, identifica-se uma abordagem implícita de cunho mecanicista, dada a ênfase na contextualização biológica e nas interações multinível na análise de dados genéticos.

Em síntese, observa-se uma diversidade de abordagens em relação à QP1. Parte dos estudos ([18, 19, 20]) se destaca por apresentar uma fundamentação teórica robusta e abrangente, contemplando dimensões formais da explicação, como o raciocínio contrastivo e mecanicista. Outros trabalhos ([37, 47]) oferecem contribuições mais periféricas, mas ainda relevantes, ao introduzirem elementos estruturantes — temporais ou biológicos — que favorecem a rastreabilidade e contextualização das inferências. Em contrapartida, estudos como [21, 26, 27] abordam a interpretabilidade de forma genérica, sem estabelecer vínculos explícitos com o arcabouço teórico da XAI. De modo geral, a literatura revisada oferece uma resposta heterogênea, porém majoritariamente positiva, à primeira pergunta de pesquisa, revelando diferentes graus de alinhamento com os referenciais conceituais mais consolidados da área.

4.2. Análise da QP2: De que formas as ontologias são utilizadas como artefatos semânticos em sistemas de XAI?

A análise das publicações selecionadas revela um amplo espectro de formas pelas quais ontologias são empregadas em sistemas de XAI, variando significativamente quanto ao nível de expressividade semântica. Os estudos de Liu et al. [24] e Kaptein et al. [42] apresentam evidências claras do uso de ontologias tanto de topo quanto específicas de domínio, utilizando-as para fundamentação de termos, estruturação causal e, em certos casos, permitindo uma integração modular que pode sustentar raciocínios contrafactuais. Essas contribuições alinham-se diretamente aos papéis previstos na QP2 e foram classificadas como evidências fortes na Tabela 1.

Outros trabalhos, como [26, 28, 29], recorrem majoritariamente a ontologias de domínio — como MedDRA [51] e Gene Ontology (GO) [52] — com o objetivo de enriquecer representações semânticas e aprimorar a interpretabilidade biológica ou clínica. Embora essas implementações fortaleçam a ancoragem semântica dos termos utilizados, raramente articulam uma estrutura mais abrangente para raciocínios causais ou contrafactuais, além de frequentemente deixarem de fazer uso de ontologias bem fundamentadas em uma ontologia de topo.

No caso de Perera and Liu [25], o uso de ontologias é apenas implícito, inserido em abordagens mais amplas baseadas em LLMs para aprendizado ontológico — uma área adjacente relevante, mas que não trata da utilização operacional de ontologias em pipelines de XAI. De forma semelhante, Li et al. [53] e Chen et al. [36] empregam embeddings e artefatos semânticos derivados de recursos estruturados (como WordNet [49], UMLS [50] e ProteinBERT [54]), contribuindo indiretamente para a fundamentação semântica, mas sem o emprego explícito de engenharia ontológica formal. Ainda que esses trabalhos não se enquadrem estritamente no escopo da XAI, a incorporação de conhecimento estruturado ao longo do pipeline de aprendizado de modelos de IA — mesmo que de forma parcial — representa um avanço rumo a sistemas semanticamente comprometidos, potencialmente mais alinhados com o mundo real e com maior capacidade de interpretação contextualizada, conforme apontaram Maass and Storey [55] e Amaral et al. [56]. Destacam-se também os estudos de Meghraoui et al. [38] e Basakin et al. [40], que apresentam abordagens neurossimbólicas integradas, demonstrando como a combinação de modelos baseados em conhecimento e aprendizado estatístico pode ser operacionalizada por meio do uso explícito de ontologias e técnicas modernas de criação de *embeddings*. Em particular, Meghraoui et al. [38] propõem uma ontologia dedicada à predição de produtividade agrícola, cujo conteúdo é formalmente avaliado e integrado a modelos de *machine learning*, ilustrando como restrições ontológicas podem sustentar raciocínios mais robustos e interpretáveis. Já os trabalhos de Zhan et al. [32] e Sung et al. [33] exploram

a ancoragem ontológica como forma de aprimorar a transparência e a contextualização das explicações, mesmo sem recorrer a ontologias de fundamentação. Em [32], a ontologia é utilizada para estruturar mecanismos explicativos em medicina de precisão, conectando entidades biomédicas a raciocínios inferenciais em sistemas de recomendação terapêutica. De forma complementar, em [33] os autores aplicam uma ontologia modular para segmentar e representar raciocínios diagnósticos, promovendo explicações mais compreensíveis em ambientes clínicos. Em ambos os casos, destaca-se o compromisso semântico estabelecido pelas ontologias no pipeline de IA, o que aproxima os modelos resultantes do raciocínio humano e do conhecimento do domínio. Por sua vez, Basakin et al. [40] enfatizam o papel de ontologias em tarefas de tomada de decisão médica, utilizando-as para representar conhecimentos biomédicos que orientam a geração de explicações simbólicas alinhadas com dados clínicos.

Por outro lado, estudos como [47, 21] empregam ontologias específicas de domínio principalmente como ferramentas para seleção de atributos ou contextualização biológica, sem recorrer à estruturação de inferências explicáveis.

A contribuição seminal de Confalonieri and Guizzardi [19] oferece uma perspectiva profunda e teoricamente fundamentada sobre a integração de ontologias em sistemas de XAI, ao apresentar uma taxonomia abrangente dos papéis que esses artefatos podem desempenhar na construção de explicações. Essa taxonomia está organizada em três dimensões principais: a primeira, voltada à modelagem de referência, destaca o uso de ontologias como base conceitual para representar de forma precisa e compartilhável o domínio de interesse, servindo como estrutura semântica para a geração de explicações alinhadas ao conhecimento especializado; a segunda, centrada no raciocínio de senso comum, explora o papel das ontologias na sustentação de inferências que extrapolam os dados disponíveis, promovendo explicações mais compreensíveis e contextualizadas para usuários não técnicos; e a terceira, relativa ao refinamento do conhecimento, trata do uso de ontologias para revisar, modular e adaptar o conteúdo explicativo, viabilizando a personalização das justificativas e o aprimoramento contínuo dos modelos explicáveis. O estudo articula com clareza como essas dimensões tornam as ontologias componentes essenciais na gestão da complexidade, na adaptação das explicações a diferentes perfis de usuários e na viabilização de raciocínios automatizados mais transparentes e robustos.

Por fim, o estudo feito em [18] leva o conceito de explicação a um novo patamar ao propor um arcabouço sistemático para explicações ontologicamente fundamentadas em sistemas de IA. O trabalho está ancorado na abordagem de *Ontology-Driven Conceptual Modeling* (ODCM), que busca explorar o papel das ontologias no suporte à construção, análise e interpretação de modelos conceituais. Embora a ODCM possa se apoiar em diferentes tipos de ontologias, o estudo em questão emprega a OntoUML — uma reformulação da linguagem UML cujos construtores sintáticos e semânticos derivam diretamente das distinções ontológicas e axiomatizações propostas pela ontologia de topo UFO (Unified Foundational Ontology) [15]. Essa integração possibilita a exploração plena de artefatos semânticos ricos, como os estereótipos ontológicos de OntoUML (e.g., *Kind*, *Role*, *Relator*), as cardinalidades fundamentadas em relações ontológicas e os chamados *ontological design patterns*, que capturam regularidades estruturais recorrentes. Tais elementos não apenas conferem maior expressividade e rigor formal aos modelos, como também servem de base para uma abordagem explicativa mais profunda.

Nesse contexto, destaca-se a técnica de *ontological unpacking*, que busca explicar descrições simbólicas — como modelos conceituais, grafos de conhecimento ou especificações formais — revelando seus compromissos ontológicos subjacentes. Isso é feito por meio da identificação de *truthmakers*, ou seja, as entidades do domínio que tornam verdadeiras as proposições modeladas (*truthbearers*). Ao tornar explícita essa camada ontológica, o processo explicativo se torna mais robusto, estruturado e cognitivamente adequado, indo além de justificativas superficiais para incorporar elementos de causalidade, dependência e temporalidade de forma logicamente coerente. Essa proposta representa um avanço significativo no uso de ontologias como suporte à geração de explicações verdadeiramente semânticas e alinhadas com os fenômenos do mundo real.

4.3. Análise da QP3: Como a fundamentação de explicações em uma ontologia afeta sua expressividade e a compreensão por parte dos usuários?

Fundamentar explicações em uma ontologia de fundamentação de forma consistente amplifica sua riqueza semântica ao tornar explícitas as relações e restrições envolvidas e melhora sua clareza e capacidade de ação por parte dos usuários finais. Liu et al. [24] demonstram que o paradigma de modelagem multinível gera explicações mais expressivas, uma vez que a estrutura em "cubo conceitual" revela dependências latentes e impõe coerência semântica, indicando maior interpretabilidade, ainda que estudos com usuários permaneçam pendentes. De forma semelhante, Zhao et al. [28] incorpora anotações da Gene Ontology ao seu visualizador explicativo *Gradient Weighted interaction Activation Mapping* (Grad-WAM), destacando sítios funcionalmente relevantes de aminoácidos e vinculando atribuições a termos estabelecidos do domínio; embora avaliações formais com usuários estejam ausentes, os vínculos com a GO oferecem um caminho claro rumo à transparência. Em Zhan et al. [32], relata-se que as explicações fundamentadas na ontologia *Myocardial infarction Ontology* (Mio) permitem inferência automatizada de conhecimento médico, tornando-as simultaneamente mais transparentes e acionáveis, enquanto Sung et al. [33] mostra que o uso da *Kyoto Encyclopedia of Genes and Genomes* (KEGG) como base ontológica em seu framework *Multi-Dimensional Transcriptomic Ruler* (MDTR) gera insights estruturados e multidimensionais sobre mecanismos de hepatotoxicidade — características que, em princípio, favorecem a compreensão do usuário ao conectar as saídas do modelo a vias biológicas bem compreendidas.

Nos casos em que avaliações empíricas foram conduzidas, os resultados confirmam o valor das explicações baseadas em ontologias. O trabalho de Smedley et al. [43] demonstra, por meio de comparações de AUC, que mascarar atributos segundo conjuntos gênicos Hallmark e GO não apenas preserva a fidelidade preditiva, como também gera explicações alinhadas ao conhecimento biológico já estabelecido, indicando maior confiança por parte dos usuários. Em Seninge et al. [45], avança-se ainda mais, mostrando que a arquitetura *Variational Autoencoder Enhanced by Gene Annotations* (VEGA) — cuja estrutura de decodificação reflete módulos gênicos definidos por especialistas — permite a construção de espaços latentes interpretáveis e gera insights acionáveis sobre a atividade de vias; feedback preliminar de usuários sugere que essas visualizações orientadas por ontologias de fato promovem maior transparência. Em contraste, Perera and Liu [25] observam que os métodos atuais baseados em LLMs (como mapas de atenção ou encadeamento de raciocínios — *Chain-of-Thought*) permanecem "longe de resolvidos" para aprendizado ontológico, ressaltando a necessidade de arcabouços semânticos mais robustos. Os trabalhos de Canito et al. [37] e Meghraoui et al. [38] reconhecem que, apesar de seus sistemas incorporarem conhecimento rico de domínio e temporalidade, ainda faltam estudos sistemáticos com usuários que avaliem a compreensão — apontando para uma lacuna relevante.

Por fim, revisões de escopo mais amplo e análises conceituais corroboram esses achados. Por exemplo, Basakin et al. [40] argumentam que explicações fundamentadas em ontologias — ao associar a semântica dos dados das saídas do modelo — tendem a promover maior confiança por parte dos usuários, mesmo que sua validação empírica ainda esteja em aberto. De forma mais aprofundada, Guizzardi and Guarino [18] propõem o uso de *ontological unpacking* como um processo explicativo que visa tornar explícitos os compromissos ontológicos embutidos em artefatos simbólicos (como modelos conceituais, grafos de conhecimento ou especificações lógicas). Essa abordagem permite revelar os *truthmakers* — entidades ontológicas que fundamentam a veracidade de uma afirmação — e sustentar explicações contrastivas, contrafactuais e logicamente coerentes, por meio da exploração sistemática dos elementos fornecidos por uma ontologia bem fundamentada, como os estereótipos ontológicos da UFO, padrões ontológicos e restrições de cardinalidade.

5. Conclusões

Embora esta pesquisa tenha seguido princípios do protocolo PRISMA 2020, trata-se de uma revisão quasi-sistemática. Isso significa que, embora tenha adotado um processo estruturado de busca, critérios de elegibilidade previamente definidos e uma análise metodológica dos achados, a revisão não possui o

caráter de exaustividade de uma revisão sistemática completa. Uma limitação relevante é que a busca foi conduzida exclusivamente na base Scopus, o que pode ter restringido a abrangência do corpus e deixado de fora trabalhos relevantes indexados em outras bases. Ainda assim, o procedimento empregado garantiu transparência, rastreabilidade e alinhamento com as boas práticas recomendadas pelo PRISMA, assegurando a consistência necessária para responder às questões de pesquisa propostas.

À luz dessas considerações, esta revisão quasi-sistemática analisou como ontologias têm sido empregadas para apoiar a Inteligência Artificial Explicável (XAI), especialmente sob perspectivas semânticas e epistêmicas. Os principais achados podem ser sumarizados a partir das perguntas de pesquisa originalmente formuladas:

QP1: Como o termo "explicação" é definido e operacionalizado na literatura atual sobre XAI?

As definições de explicação ainda são heterogêneas, com enfoques distintos em dimensões *mecanicistas* (como entradas afetam saídas), *contrastivas* (explicações "por que A e não B?"), e *sociais* (considerando o perfil e contexto do usuário). Contudo, poucos trabalhos apresentam abordagens sistematicamente fundamentadas. Apenas uma minoria, como Guizzardi and Guarino [18] e Confalonieri and Guizzardi [19], propõe modelos explicativos com base em teorias filosóficas ou ontológicas robustas.

QP2: De que formas as ontologias são utilizadas como artefatos semânticos em sistemas de XAI?

As ontologias são majoritariamente utilizadas para a *fundamentação de termos* e a *contextualização semântica*, especialmente em domínios biomédicos. Ontologias específicas de domínio, como Gene Ontology e KEGG, aparecem com mais frequência, mas seu uso tende a ser superficial. Ontologias de fundamentação, como a UFO, são raramente exploradas, apesar de seu potencial para apoiar *raciocínio causal*, *explicações contrafactuais* e garantir menor ambiguidade semântica. Trabalhos como de Guizzardi and Guarino [18] demonstram que o uso de modelos conceituais ontologicamente bem fundados permite maior expressividade e clareza nos comprometimentos semânticos das explicações.

QP3: Como a fundamentação de explicações em ontologias afeta sua expressividade e a compreensão por parte dos usuários?

A fundamentação ontológica tende a ampliar significativamente a *expressividade semântica* das explicações, tornando explícitas as relações e restrições entre conceitos. Visualizações baseadas em ontologias, como as que utilizam a Gene Ontology ou o KEGG, permitem associações mais ricas e estruturadas. Há indícios de que essas abordagens aumentam a *transparência percebida* e a *confiança* dos usuários. No entanto, a maior parte dos estudos carece de avaliações empíricas sistemáticas sobre a compreensão e a utilidade das explicações.

Com base nesses achados, delineiam-se algumas recomendações para pesquisas futuras. É necessário promover a **criação de ontologias de domínio fundamentadas em ontologias de fundamentação**, explorando distinções ontológicas bem estabelecidas para estruturar explicações contrastivas, causais e temporais. Ainda, é importante ampliar **estudos sistemáticos com usuários** que avaliem a compreensão, confiança e eficácia das explicações baseadas em ontologias; outras questões ainda em aberto incluem desenvolver **métricas e benchmarks padronizados** que possibilitem a avaliação comparativa de abordagens explicativas guiadas por semântica estruturada, ampliar a investigação sobre **integração entre pipelines subsimbólicos e estruturas simbólicas formais por meio da IA neurossimbólica**, com vistas a garantir maior robustez estatística e inteligibilidade conceitual.

A consolidação da XAI como uma disciplina verdadeiramente transparente, confiável e centrada no ser humano dependerá da combinação entre avanços técnicos e fundamentos semânticos sólidos, possibilitados pelo uso criterioso de ontologias bem estruturadas.

Declaration on Generative AI

In the preparation of this manuscript, the authors made limited use of the ChatGPT-4o generative AI model. The tool was employed exclusively for linguistic assistance, including rephrasing, improving readability and style, and checking grammar and spelling, as outlined in the CEUR-WS GenAI Usage Taxonomy. No part of the research design, analysis, or interpretation of results was carried out by AI. All suggested edits were critically reviewed, adapted, and validated by the authors, who take full responsibility for the final content of this publication.

Referências

[1] M. J. Page, J. E. McKenzie, P. M. Bossuyt, et al., The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, Bmj (2021) n71. doi:10.1136/bmj.n71.

[2] H. Sheikh, C. Prins, E. Schrijvers, Artificial Intelligence: Definition and Background, Springer International Publishing, Cham, 2023, pp. 15–41. doi:10.1007/978-3-031-21448-6_2.

[3] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, et al., Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, Information Fusion 58 (2020) 82–115. doi:10.1016/j.inffus.2019.12.012.

[4] R. Studer, V. Benjamins, D. Fensel, Knowledge engineering: Principles and methods, Data & Knowledge Engineering 25 (1998) 161–197. doi:10.1016/s0169-023x(97)00056-6.

[5] W. Maass, V. C. Storey, Pairing conceptual modeling with machine learning, Data amp; Knowledge Engineering 134 (2021) 101909. doi:10.1016/j.datak.2021.101909.

[6] S. Ali, T. Abuhmed, S. El-Sappagh, et al., Explainable artificial intelligence (xai): What we know and what is left to attain trustworthy artificial intelligence, Information Fusion 99 (2023) 101805. doi:https://doi.org/10.1016/j.inffus.2023.101805.

[7] E. Rajabi, K. Etminani, Knowledge-graph-based explainable ai: A systematic review, Journal of Information Science 50 (2022) 1019–1029. URL: http://dx.doi.org/10.1177/01655515221112844. doi:10.1177/01655515221112844.

[8] R. Confalonieri, L. Coba, B. Wagner, et al., A historical perspective of explainable Artificial Intelligence, WIREs Data Mining and Knowledge Discovery 11 (2021) e1391. doi:10.1002/widm.1391.

[9] B. Mittelstadt, C. Russell, S. Wachter, Explaining explanations in ai, in: Proceedings of the Conference on Fairness, Accountability, and Transparency, Fat* '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 279–288. doi:10.1145/3287560.3287574.

[10] M. T. Ribeiro, S. Singh, C. Guestrin, "why should I trust you?": Explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016, 2016, pp. 1135–1144.

[11] S. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, 2017. doi:10.48550/arxiv.1705.07874.

[12] G. Schwalbe, B. Finzel, A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts, Data Mining and Knowledge Discovery 38 (2024) 3043–3101. doi:10.1007/s10618-022-00867-8.

[13] 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.

[14] S. Badreddine, A. d'Avila Garcez, L. Serafini, et al., Logic Tensor Networks, Artificial Intelligence 303 (2022) 103649. doi:10.1016/j.artint.2021.103649.

[15] G. Guizzardi, A. Botti Benevides, C. M. Fonseca, et al., UFO: Unified Foundational Ontology, Applied Ontology 17 (2022) 167–210. doi:10.3233/ao-210256.

[16] S. Borgo, R. Ferrario, A. Gangemi, et al., Dolce: A descriptive ontology for linguistic and cognitive engineering, Applied Ontology 17 (2022) 45–69. doi:10.3233/ao-210259.

[17] J. N. Otte, J. Beverley, A. Ruttenberg, Bfo: Basic formal ontology, Applied Ontology 17 (2022) 17–43. doi:10.3233/ao-220262.
[18] G. Guizzardi, N. Guarino, Explanation, semantics, and ontology, Data Knowledge Engineering 153 (2024) 102325. doi:https://doi.org/10.1016/j.datak.2024.102325.
[19] R. Confalonieri, G. Guizzardi, On the multiple roles of ontologies in explainable ai, 2023. doi:10.48550/arxiv.2311.04778.
[20] S. Chari, D. M. Gruen, O. Seneviratne, et al., Directions for explainable knowledge-enabled systems, 2020. doi:10.48550/arxiv.2003.07523.
[21] P. Sun, Y. Wu, C. Yin, et al., Molecular Subtyping Of Cancer Based On Distinguishing Co-expression Modules And Machine Learning, Front Genet 13 (2022). doi:10.3389/fgene.2022.866005.
[22] C. Lu, S. Pathak, G. Englebienne, et al., Channel Contribution In Deep Learning Based Automatic Sleep Scoring - How Many Channels Do We Need?, Ieee Trans Neural Syst Rehabil Eng 31 (2023). doi:10.1109/tnsre.2022.3227040.
[23] Y. Li, Y. Liu, J. Hou, et al., A Center-anchored Adaptive Hierarchical Graph Neural Network With Application In Structure-aware Recognition Of Enzyme Catalytic Specificity, Neurocomputing 619 (2025). doi:10.1016/j.neucom.2024.129155.
[24] Z. Liu, Z. Zhang, X. Zeng, et al., Representation And Association Of Chinese Financial Equity Knowledge Driven By Multilayer Ontology, Data Inf Manag 6 (2022) 3.0. doi:10.1016/j.dim.2022.100009.
[25] O. Perera, J. Liu, Exploring Large Language Models For Ontology Learning, Issue Inf Syst 25 (2024) 4.0. doi:10.48009/4_iis_2024_124.
[26] X. Ma, T. Wu, G. Li, et al., Dse-hngcn: Predicting the frequencies of drug-side effects based on heterogeneous networks with mining interactions between drugs and side effects, Journal of Molecular Biology 437 (2025) 168916. doi:https://doi.org/10.1016/j.jmb.2024.168916, emerging Artificial Intelligence Methodologies in Computational Biology.
[27] Y. Wei, T. Zhang, B. Wang, et al., Indelpred: Improving The Prediction And Interpretation Of Indel Pathogenicity Within The Clinical Genome, Hum Genet Genom Adv 5 (2024) 4.0. doi:10.1016/j.xhgg.2024.100325.
[28] S. Zhao, Z. Cui, G. Zhang, et al., Mgppi: Multiscale Graph Neural Networks For Explainable Protein–protein Interaction Prediction, Front Genet 15 (2024). doi:10.3389/fgene.2024.1440448.
[29] L. Rasbach, A. Caliskan, F. Saderi, et al., An Orchestra Of Machine Learning Methods Reveals Landmarks In Single-cell Data Exemplified With Aging Fibroblasts, Plos One 19 (2024) 4.0. doi:10.1371/journal.pone.0302045.
[30] S. Dubey, S. Singh, D. Verma, et al., Genocare Prognosticator Model: Host Genetics Predict Severity Of Infectious Disease, Scalable Comput Pract Exp 26 (2025) 2.0. doi:10.12694/scpe.v26i2.3774.
[31] Y. Zhao, J. Shao, Y. Asmann, Assessment And Optimization Of Explainable Machine Learning Models Applied To Transcriptomic Data, Genomics Proteomics Bioinform 20 (2022) 5.0. doi:10.1016/j.gpb.2022.07.003.
[32] C. Zhan, S. Ren, Y. Zhang, et al., Mio: An Ontology For Annotating And Integrating Medical Knowledge In Myocardial Infarction To Enhance Clinical Decision Making, Comput Biol Med 190 (2025). doi:10.1016/j.compbiomed.2025.110107.
[33] I. Sung, S. Lee, D. Bang, et al., Mdtr: A Knowledge-guided Interpretable Representation For Quantifying Liver Toxicity At Transcriptomic Level, Front Pharmacol 15 (2024). doi:10.3389/fphar.2024.1398370.
[34] S. Li, F. Dong, R. Li, et al., A Dual Medical Ontology And Relational Graph Framework With Neural Ordinary Differential Equations For Diagnostic Prediction, Neurocomputing 647 (2025). doi:10.1016/j.neucom.2025.130519.
[35] X. Zheng, D. Meng, D. Chen, et al., Sccat: An Explainable Capsulating Architecture For Sepsis Diagnosis Transferring From Single-cell Rna Sequencing, Plos Comput Biol 20 (2024) 10.0. doi:10.1371/journal.pcbi.1012083.
[36] Z. Chen, Z. He, X. Liu, et al., Evaluating Semantic Relations In Neural Word Embeddings With

Biomedical And General Domain Knowledge Bases, Bmc Med Informatics Decis Mak 18 (2018). doi:10.1186/s12911-018-0630-x.

[37] A. Canito, J. Corchado, G. Marreiros, A Systematic Review On Time-constrained Ontology Evolution In Predictive Maintenance, Artif Inl Rev 55 (2022) 4.0. doi:10.1007/s10462-021-10079-z.

[38] K. Meghraoui, T. Racharak, K. Ait, et al., A New Integrated Neurosymbolic Approach For Crop-yield Prediction Using Environmental Data And Salite Imagery At Field Scale, Artif Inl Geosci 6 (2025) 1.0. doi:10.1016/j.aiig.2025.100125.

[39] C. Guan, A. Gong, Y. Zhao, et al., Interpretable Machine Learning Model For New-onset Atrial Fibrillation Prediction In Critically Ill Patients: A Multi-center Study, Crit Care 28 (2024) 1.0. doi:10.1186/s13054-024-05138-0.

[40] A. Basakin, Y. Kulchin, V. Gribova, et al., Prospects For The Development Of Laser-based Directed Energy Deposition Additive Process Based On Ai Technologies, Bull Russ Acad Sci Phys 88 (2024) 3.0. doi:10.1134/s1062873824709711.

[41] Q. Li, Y. Yu, P. Kossinna, et al., Xa4c: Explainable Representation Learning Via Autoencoders Revealing Critical Genes, Plos Comput Biol 19 (2023) 10.0. doi:10.1371/journal.pcbi.1011476.

[42] F. Kaptein, B. Kiefer, A. Cully, et al., A Cloud-based Robot System For Long-term Interaction: Principles, Implementation, Lessons Learned, Acm Trans Hum Robot Interact 11 (2022) 1.0. doi:10.1145/3481585.

[43] N. Smedley, D. Aberle, W. Hsu, Using Deep Neural Networks And Interpretability Methods To Identify Gene Expression Patterns That Predict Radiomic Features And Histology In Non-small Cell Lung Cancer, J Med Imaging 8 (2021) 3.0. doi:10.1117/1.jmi.8.3.031906.

[44] C. Ruiz-Arenas, I. Mar, L. Wang, et al., Netactivity Enhances Transcriptional Signals By Combining Gene Expression Into Robust Gene Set Activity Scores Through Interpretable Autoencoders, Nucleic Acids Res 52 (2024) 9.0. doi:10.1093/nar/gkae197.

[45] L. Seninge, I. Anastopoulos, H. Ding, et al., Vega Is An Interpretable Generative Model For Inferring Biological Network Activity In Single-cell Transcriptomics, Nat Commun 12 (2021) 1.0. doi:10.1038/s41467-021-26017-0.

[46] M. Singha, L. Pu, G. Srivastava, et al., Unlocking The Potential Of Kinase Targets In Cancer: Insights From Canceromicsnet, An Ai-driven Approach To Drug Response Prediction In Cancer, Cancers 15 (2023) 16.0. doi:10.3390/cancers15164050.

[47] K. Askland, D. Strong, M. Wright, et al., The Translational Machine: A Novel Machine-learning Approach To Illuminate Complex Genetic Architectures, Genet Epidemiol 45 (2021) 5.0. doi:10.1002/gepi.22383.

[48] H. Townsend, K. Rosenberger, L. Vanderlinden, et al., Evaluating Methods For Integrating Single-cell Data And Genetics To Understand Inflammatory Disease Complexity, Front Immunol 15 (2024). doi:10.3389/fimmu.2024.1454263.

[49] G. A. Miller, WordNet: A lexical database for English, in: Human Language Technology: Proceedings of a Workshop held at Plainsboro, New Jersey, March 8-11, 1994, 1994.

[50] O. Bodenreider, The Unified Medical Language System (UMLS): integrating biomedical terminology, Nucleic Acids Research 32 (2004) 267d–270. doi:10.1093/nar/gkh061.

[51] E. G. Brown, L. Wood, S. Wood, The medical dictionary for regulatory activities (meddra), Drug Safety 20 (1999) 109–117. doi:10.2165/00002018-199920020-00002.

[52] M. Ashburner, C. A. Ball, J. A. Blake, et al., Gene ontology: tool for the unification of biology, Nature Genetics 25 (2000) 25–29. doi:10.1038/75556.

[53] Y. Li, Y. Liu, J. Hou, et al., A center-anchored adaptive hierarchical graph neural network with application in structure-aware recognition of enzyme catalytic specificity, Neurocomputing 619 (2025) 129155. doi:https://doi.org/10.1016/j.neucom.2024.129155.

[54] N. Brandes, D. Ofer, Y. Peleg, et al., Proteinbert: a universal deep-learning model of protein sequence and function, Bioinformatics 38 (2022) 2102–2110. doi:10.1093/bioinformatics/btac020.

[55] W. Maass, V. C. Storey, Pairing conceptual modeling with machine learning, Data Knowl. Eng. 134 (2021) 101909. doi:10.1016/j.datak.2021.101909.

[56] G. Amaral, F. Baião, G. Guizzardi, Foundational ontologies, ontology-driven conceptual modeling,

and their multiple benefits to data mining, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 11 (2021) e1408.