

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/386534303>

METODOLOGIAS SOBRE APRENDIZAGEM DE ONTOLOGIAS: UMA REVISÃO SISTEMÁTICA DE LITERATURA

Article in *Revista Científica Multidisciplinar Núcleo do Conhecimento* · December 2024

DOI: 10.32749/nucleodoconhecimento.com.br/tecnologia/aprendizagem-de-

CITATIONS

0

READS

22

2 authors:



Gercina Angela de Lima

Federal University of Minas Gerais

142 PUBLICATIONS 560 CITATIONS

[SEE PROFILE](#)



Patrícia Lopes

Serviço Federal de Processamento de Dados

2 PUBLICATIONS 3 CITATIONS

[SEE PROFILE](#)



METODOLOGIAS SOBRE APRENDIZAGEM DE ONTOLOGIAS: UMA REVISÃO SISTEMÁTICA DE LITERATURA

ARTIGO DE REVISÃO

LOPES, Patrícia¹, LIMA, Gercina Ângela de²

LOPES, Patrícia. LIMA, Gercina Ângela de. **Metodologias sobre aprendizagem de ontologias: uma revisão sistemática de literatura.** Revista Científica Multidisciplinar Núcleo do Conhecimento. Ano 09, Ed. 11, Vol. 10, pp. 139-171. Novembro de 2024. ISSN: 2448-0959, Link de acesso: <https://www.nucleodoconhecimento.com.br/tecnologia/aprendizagem-de-ontologias>, DOI: 10.32749/nucleodoconhecimento.com.br/tecnologia/aprendizagem-de-ontologias

RESUMO

Contextualização: Ontologias vêm ganhando popularidade e reconhecimento em diferentes contextos. Entretanto, com o aumento exponencial de dados não estruturados na Web, o processo de construção de ontologias tornou-se complexo, demorado e dispendioso. Surgem, então, as abordagens de geração automática ou semiautomática de ontologias. Dado o contexto atual de transformação digital pelo qual passam diversos países, incluindo o Brasil, entende-se que as pesquisas sobre a automatização da geração de ontologias podem contribuir para a exploração e melhor aplicação desse tipo de Sistema de Organização do Conhecimento (SOC) no âmbito do governo eletrônico. **Objetivo:** Este estudo apresenta uma Revisão Sistemática da Literatura (RSL) sobre a geração automática e semiautomática de ontologias, com o objetivo de evidenciar o estágio atual da pesquisa nessa área. **Métodos:** Adotou-se o método proposto por Kitchenham (2004), que descreve procedimentos para realizar revisões sistemáticas de literatura, considerando a formulação de perguntas de pesquisa, o desenvolvimento de um protocolo, a identificação e seleção de estudos, avaliação de qualidade, extração e síntese de dados e relato dos resultados. **Resultados:** A revisão revelou uma diversidade significativa de abordagens e metodologias sendo exploradas para a geração automática e semiautomática de ontologias. No entanto muitas metodologias ainda enfrentam desafios, principalmente no que diz respeito ao detalhamento de fases e etapas que sejam facilmente replicáveis. **Conclusões:** Futuras pesquisas devem focar na criação de metodologias mais detalhadas e sistemáticas para aumentar a replicabilidade e a aplicação prática. As descobertas deste estudo podem inspirar investigações adicionais e estimular a implementação das ontologias no âmbito no



governo eletrônico, com vistas a aprimorar a eficiência, a transparência e a acessibilidade dos serviços públicos digitais.

Palavras-chave: Ciência da Informação, Ontologias, Aprendizagem de ontologias, Governo eletrônico.

1. INTRODUÇÃO

A Ciência da Informação tem passado por transformações significativas, especialmente na organização e representação do conhecimento, em que as ontologias emergem como ferramentas cruciais para a estruturação e interpretação de informações complexas. Contudo a explosão de dados não estruturados na Web impõe desafios ao processo tradicional de construção de ontologias, tornando-o complexo, demorado e dispendioso.

Apesar do crescente interesse na construção de ontologias, sua utilização em larga escala ainda é limitada devido às demandas consideráveis de tempo, custo e recursos envolvidos no processo. Para contribuir com o avanço da investigação nesse campo, diversas pesquisas têm se concentrado na automatização de diferentes etapas da construção de ontologias.

Este estudo apresenta uma Revisão Sistemática da Literatura (RSL) sobre a geração automática e semiautomática de ontologias[3], área conhecida como aprendizagem de ontologias ou *ontology learning*, buscando agregar conhecimento sobre as abordagens, metodologias e ferramentas mais recentes. A revisão de artigos científicos visa identificar metodologias, etapas, vantagens e lacunas apontadas pelos autores, bem como os principais resultados. Compreender a dinâmica e o estado atual da pesquisa nesse campo do conhecimento contribuirá para o desenvolvimento e o aprimoramento dessa área crucial na representação do conhecimento.

Para a organização e gestão das referências, utilizou-se o *software* Mendeley, enquanto o Atlas TI foi empregado para a análise qualitativa dos estudos selecionados. A pesquisa visou identificar trabalhos que aplicaram métodos de aprendizagem de ontologias entre 2018 e 2023, abordando questões como a identificação de metodologias específicas, etapas envolvidas, vantagens, diferenciais



e lacunas existentes. As bases de dados consultadas incluem *ACM Digital Library*, *IEEE Xplore Digital Library*, *Scopus Preview* e *Web of Science*, utilizando termos de busca em inglês e português para garantir a abrangência e a relevância dos resultados.

O protocolo da RSL definiu critérios de inclusão e exclusão para assegurar a qualidade e a pertinência dos estudos analisados. Foram considerados apenas trabalhos com relação direta à temática de aprendizagem de ontologias, disponíveis integralmente e publicados em inglês, português ou espanhol. A exclusão de duplicatas e a filtragem de trabalhos irrelevantes resultaram em 183 textos, reduzidos a 170 após a verificação de disponibilidade do texto completo. A análise final concentrou-se em 22 estudos, selecionados pela sua relevância e aderência aos critérios estabelecidos.

Este estudo contextualiza a importância da geração automática de ontologias, destacando sua aplicação prática e potencial de inovação no campo do governo eletrônico. A revisão detalhada das metodologias existentes permite identificar avanços e desafios, contribuindo para o desenvolvimento de abordagens mais eficazes e robustas. A clareza metodológica e a replicabilidade são aspectos fundamentais para a evolução dessa área, proporcionando uma base sólida para futuras pesquisas e implementações práticas. O artigo está organizado da seguinte forma: esta seção 1, que é introdutória; a seção 2, que detalha os procedimentos metodológicos adotados na RSL; a seção 3, que apresenta os resultados e discute as principais metodologias identificadas, incluindo diferenciais, lacunas e resultados; e a seção 4, que traz as considerações finais e sugestões para pesquisas futuras.

2. PROCEDIMENTOS METODOLÓGICOS ADOTADOS NA REVISÃO

Os procedimentos metodológicos para a Revisão Sistemática da Literatura foram respaldados no método proposto por Kitchenham (2004). Para organização e gestão das referências, foi utilizado o *software* Mendeley. O *software* Atlas TI foi utilizado para análise qualitativa do material selecionado. As pesquisas referenciadas foram selecionadas considerando maior aproximação dos resultados apresentados com a



temática desta pesquisa e estão organizadas nas categorias de análise previamente identificadas no protocolo da revisão, descrito no Quadro 1, a seguir.

Quadro 1 – Resumo Protocolo da Revisão Sistemática

Temática Principal	Geração Automática de Ontologias
Áreas de Conhecimento	Ciência da Informação, Ciência da Computação.
Protocolo de Revisão Sistemática da Literatura (PRSL)	Apresenta a estrutura metodológica para a execução da Revisão da Literatura sobre métodos para o processo de aprendizagem de ontologias.
Pergunta Principal do Estudo	Quais trabalhos aplicaram métodos de aprendizagem de ontologias entre os anos de 2018 e 2023?
Perguntas Secundárias	(1) É possível identificar a utilização de alguma metodologia de aprendizagem de ontologias?
	(2) Quais as etapas/fases da metodologia ora adotada?
	(3) Quais as vantagens e os diferenciais da metodologia adotada?
	(4) Quais as lacunas existentes nessas metodologias?
	(5) Quais os principais resultados evidenciados?

Fonte: elaborado pelas autoras (2024).

Partindo-se das perguntas de pesquisa, seguiu-se para o levantamento dos termos de busca, considerando inicialmente as variações nos idiomas inglês e português, conforme disposto no Quadro 2, a seguir.

Quadro 2 – Termos de busca

TERMOS EM INGLÊS	TERMOS EM PORTUGUÊS
[Ontology Learning] OR [Ontologies Learning]	[Aprendizagem de Ontologia] OR [Aprendizagem de Ontologias]
[Ontology Engineering] OR [Ontologies Engineering]	[Engenharia de ontologia] OR [Engenharia de ontologias]
[Ontology Extraction] OR [Ontologies Extraction]	[Extração de Ontologia] OR [Extração de Ontologias]
[Automatic ontology Generation] OR [Automatic ontologies Generation]	[Geração automática de ontologia] OR [Geração automática de ontologias]



Fonte: elaborado pelas autoras (2024).

Após a definição dos termos de busca, considerando a aderência à temática da pesquisa, foi realizado o levantamento das bases de dados, chegando-se às seguintes: *ACM Digital Library*; *IEEE Xplore Digital Library*; *Scopus Preview* e *Web of Science*. Inicialmente, executaram-se as expressões de busca indicadas no Quadro 3 em cada uma das bases de dados. A partir da lista de estudos retornados, os resultados foram exportados em formato RIS para importação na ferramenta Mendeley.

Quadro 3 – *Strings* de busca

BASE DE DADOS		STRING UTILIZADA
ACM Library	Digital	[Title: "ontology learning"] OR [Title: "ontologies learning"] OR [Title: "aprendizagem de ontologia"] OR [Title: "aprendizagem de ontologias"] OR [Title: "ontology engineering"] OR [Title: "ontologies engineering"] OR [Title: "engenharia de ontologia"] OR [Title: "engenharia de ontologias"] OR [Title: "ontology extraction"] OR [Title: "ontologies extraction"] OR [Title: "extração de ontologia"] OR [Title: "extração de ontologias"] OR [Title: "automatic ontology generation"] OR [Title: "automatic ontologies generation"] OR [Title: "geração automática de ontologia"] OR [Title: "geração automática de ontologias"] OR [Abstract: "ontology learning"] OR [Abstract: "ontologies learning"] OR [Abstract: "aprendizagem de ontologia"] OR [Abstract: "aprendizagem de ontologias"] OR [Abstract: "ontology engineering"] OR [Abstract: "ontologies engineering"] OR [Abstract: "engenharia de ontologia"] OR [Abstract: "engenharia de ontologias"] OR [Abstract: "ontology extraction"] OR [Abstract: "ontologies extraction"] OR [Abstract: "extração de ontologia"] OR [Abstract: "extração de ontologias"] OR [Abstract: "automatic ontology generation"] OR [Abstract: "automatic ontologies generation"] OR [Abstract: "geração automática de ontologia"] OR [Abstract: "geração automática de ontologias"] OR [Keywords: "ontology learning"] OR [Keywords: "ontologies learning"] OR [Keywords: "aprendizagem de ontologia"] OR [Keywords: "aprendizagem de ontologias"] OR [Keywords: "ontology engineering"] OR [Keywords: "ontologies engineering"] OR [Keywords: "engenharia de ontologia"] OR [Keywords: "engenharia de ontologias"] OR [Keywords: "ontology extraction"] OR [Keywords: "ontologies extraction"] OR [Keywords: "extração de ontologia"] OR [Keywords: "extração de ontologias"] OR [Keywords: "automatic ontology generation"] OR [Keywords: "automatic ontologies generation"] OR [Keywords: "geração automática de ontologia"] OR [Keywords: "geração automática de ontologias"] AND [E-Publication Date: (01/01/2018 TO 12/31/2023)]
IEEE		("Document Title": "Ontology Learning") OR ("Document Title": "Ontology Learning") OR "Document Title": "Ontologies Learning" OR "Document Title": "Aprendizagem de Ontologia" OR "Document Title": "Aprendizagem de Ontologias" OR "Document Title": "Ontology Engineering" OR "Document Title": "Ontologies Engineering" OR "Document Title": "Engenharia de ontologia" OR "Document Title": "Engenharia de ontologies" OR "Document Title": "Ontology Extraction" OR "Document Title": "Ontologies Extraction" OR



	"Document Title": "Extração de Ontologia" OR "Document Title": "Extração de Ontologias" OR "Document Title": "Automatic ontology Generation" OR "Document Title": "Automatic ontologies Generation" OR "Document Title": "Geração automática de ontologia" OR "Document Title": "Geração automática de ontologies") OR ("Abstract": "Ontology Learning" OR "Abstract": "Ontologies Learning" OR "Abstract": "Aprendizagem de Ontologia" OR "Abstract": "Aprendizagem de Ontologias" OR "Abstract": "Ontology Engineering" OR "Abstract": "Ontologies Engineering" OR "Abstract": "Engenharia de ontologia" OR "Abstract": "Engenharia de ontologies" OR "Abstract": "Ontology Extraction" OR "Abstract": "Ontologies Extraction" OR "Abstract": "Extração de Ontologia" OR "Abstract": "Extração de Ontologias" OR "Abstract": "Automatic ontology Generation" OR "Abstract": "Automatic ontologies Generation" OR "Abstract": "Geração automática de ontologia" OR "Abstract": "Geração automática de ontologies") OR ("Author Keywords": "Ontology Learning" OR "Author Keywords": "Ontologies Learning" OR "Author Keywords": "Aprendizagem de Ontologia" OR "Author Keywords": "Aprendizagem de Ontologias" OR "Author Keywords": "Ontology Engineering" OR "Author Keywords": "Ontologies Engineering" OR "Author Keywords": "Engenharia de ontologia" OR "Author Keywords": "Engenharia de ontologies" OR "Author Keywords": "Ontology Extraction" OR "Author Keywords": "Ontologies Extraction" OR "Author Keywords": "Extração de Ontologia" OR "Author Keywords": "Extração de Ontologias" OR "Author Keywords": "Automatic ontology Generation" OR "Author Keywords": "Automatic ontologies Generation" OR "Author Keywords": "Geração automática de ontologia" OR "Author Keywords": "Geração automática de ontologies") Filters Applied: JournalsBooks2018 – 2023
Scopus	TITLE-ABS-KEY ("Ontology Learning" OR "Ontologies Learning" OR "Aprendizagem de Ontologia" OR "Aprendizagem de Ontologias" OR "Ontology Engineering" OR "Ontologies Engineering" OR "Engenharia de ontologia" OR "Engenharia de ontologies" OR "Ontology Extraction" OR "Ontologies Extraction" OR "Extração de Ontologia" OR "Extração de Ontologias" OR "Automatic ontology Generation" OR "Automatic ontologies Generation" OR "Geração automática de ontologia" OR "Geração automática de ontologies") AND PUBYEAR > 2017 AND PUBYEAR < 2024
Web of Science	"Ontology Learning" OR "Ontologies Learning" OR "Aprendizagem de Ontologia" OR "Aprendizagem de Ontologias" OR "Ontology Engineering" OR "Ontologies Engineering" OR "Engenharia de ontologia" OR "Engenharia de ontologies" OR "Ontology Extraction" OR "Ontologies Extraction" OR "Extração de Ontologia" OR "Extração de Ontologias" OR "Automatic ontology Generation" OR "Automatic ontologies Generation" OR "Geração automática de ontologia" OR "Geração automática de ontologies" P.S. Utilizou-se como parâmetro de busca o item "Topic", que realiza a busca nas seções Título, Resumo e Palavras-chave. Utilizou-se o filtro Ano de publicação, adotando-se o intervalo entre 2018 e 2023.

Fonte: elaborado pelas autoras (2024).

Na primeira etapa de seleção dos textos, foram excluídas as duplicatas, resultando-se em 183 textos. Considerando a exclusão dos trabalhos que não disponibilizavam o



texto completo, chegou-se a um total de 170 textos. Essas etapas foram realizadas entre os meses de setembro e outubro de 2023.

Em seguida, foram identificados os trabalhos aderentes aos critérios de inclusão e exclusão, indicados nos Quadros 4 e 5, refletindo o *corpus* final para a etapa de análise, chegando-se a 22 textos. Essa etapa foi realizada entre novembro de 2023 e janeiro de 2024.

Quadro 4 – Critérios de Inclusão

CRITÉRIO	DESCRIÇÃO DO CRITÉRIO DE INCLUSÃO
CI1	Título possui relação com a temática aprendizagem de ontologias.
CI2	Resumo possui relação com a temática aprendizagem de ontologias.
CI3	Serão incluídos trabalhos publicados e disponíveis integralmente nas bases de dados científicas buscadas.
CI4	Serão incluídos trabalhos que tenham aplicado um método de geração automática ou semiautomática de ontologias.
CI5	Trabalhos nos idiomas Inglês, Português e Espanhol.
CI6	Período de publicação compreendido entre 2018 e 2023.
CI7	Serão considerados trabalhos dos seguintes tipos: artigo de periódico, capítulo de livro, artigo de evento.

Fonte: elaborado pelas autoras (2024).

Quadro 5 – Critérios de Exclusão

CRITÉRIO	DESCRIÇÃO DO CRITÉRIO DE EXCLUSÃO
CE1	Título não possui relação com a temática aprendizagem de ontologias.
CE2	Resumo não possui relação com a temática aprendizagem de ontologias.
CE3	Serão excluídos trabalhos que não apresentem resumo/abstract.
CE4	Serão excluídos <i>short papers</i> e trabalhos disponibilizados em bases preprint.
CE5	Serão excluídos documentos que não estejam em português, inglês ou espanhol.
CE6	Serão excluídos trabalhos que não estejam disponíveis em texto completo.
CE7	Serão excluídos os trabalhos que não apliquem o processo de geração automática ou semiautomática de ontologias.

Fonte: elaborado pelas autoras (2024).



Finalizada a etapa de extração dos dados, foram realizadas a análise, a interpretação e a documentação dos resultados encontrados nos estudos, tendo em vista a pergunta principal e as perguntas secundárias previstas nesta RSL. Essa etapa foi realizada entre os meses de janeiro e fevereiro de 2024 com o apoio do *software* Atlas TI.

3. RESULTADOS E DISCUSSÃO

A construção de ontologias é crucial na Ciência da Informação para a estruturação e interpretação de informações complexas. Tem sido objeto de intensas pesquisas e inovações, abrangendo diversas áreas e aplicando uma variedade de abordagens para melhorar a eficiência e a precisão na estruturação do conhecimento. A seguir, no Quadro 6, um resumo das metodologias identificadas nos estudos referenciados:

Quadro 6 – Panorama das metodologias identificadas nos estudos

Título do artigo	Referência	Descrição da Metodologia
Catálogos de Ontologias para IoT e Cidades Inteligentes	Gyrard, Zimmermann e Sheth, 2018	Análise e aprimoramento de catálogos ontológicos, destacando a utilidade e atualizações contínuas do catálogo LOV4IoT.
Previsão de Pontuações de Axiomas	Malchiodi e Tettamanzi, 2018; Malchiodi, Da Costa Pereira e Tettamanzi, 2018	Utilização de algoritmos de suporte vetorial e técnicas de inferência para prever pontuações de axiomas OWL, mostrando eficácia em comparação a métodos tradicionais.
Aprendizado de Ontologia em Dados Taxonômicos	Deb <i>et al.</i> , 2018	Extração de termos, conceitos e relações de textos taxonômicos no sistema agrícola, abordando limitações dos sistemas existentes.
Extração de Conhecimento de Documentos HTML	Rodriguez e Aguilar, 2018	Aplicação da PLN para converter texto de documentos HTML em conhecimento estruturado, utilizando grafos de aprendizagem.
LAM (Modulador de Algoritmo de Aprendizado)	Enea <i>et al.</i> , 2018	Combinação de extração automática de triplas com validação humana, utilizando SPARQL para consultas hipotéticas e visualizações de incertezas na ontologia.



Aprendizagem de Ontologias de Bancos de Dados Relacionais Heterogêneos	Aggoune, 2018	Utilização de WordNet e medidas de similaridade semântica para gerar ontologias no contexto de riscos alimentares, apresentando processos sequenciais detalhados e comparações com outras ferramentas.
Plataforma EasyOnto	Yadav <i>et al.</i> , 2018	Promoção da criação colaborativa de ontologias semiformal por meio de uma interface gráfica simples, facilitando a participação em massa e o desenvolvimento ágil de ontologias.
Evolução Gramatical e Teoria da Possibilidade	Nguyen e Tettamanzi, 2019	Descoberta de axiomas disjuntos em ontologias a partir de dados RDF, mostrando precisão e abrangência satisfatórias.
KE4WoT	Noura <i>et al.</i> , 2019	Foco na extração automatizada de conhecimento para aplicações na Web Semântica das Coisas (WoT), com avaliações mostrando eficácia na identificação de conceitos pertinentes.
Framework BOLT-K	Vedula, Maneriker e Parthasarathy, 2019	Base em LSTM, adaptação de conhecimento de domínios relacionados para o aprendizado de ontologias, com avaliações comparativas e representações gráficas detalhadas.
Computer Science Ontology (CSO)	Salatino <i>et al.</i> , 2020	Construção de ontologia automaticamente para áreas de pesquisa em Ciência da Computação, empregando aprendizado de máquina e PLN, com ferramentas associadas para classificação de artigos e exploração da ontologia.
Extração de Estruturas Ontológicas em Língua Espanhola	Gómez-Suta, Echeverry-Correa e Soto-Mejía, 2020	Extração de conceitos e relações em textos sobre o conflito armado colombiano, utilizando técnicas de agrupamento semântico e avaliação de estruturas extraídas.
Reutilização de Ontologias	Alharbi, Tamma e Grasso, 2021	Investigação da prática <i>versus</i> teoria na reutilização de ontologias, com <i>insights</i> valiosos de desenvolvedores e obstáculos identificados.



Expansão Automática de Ontologias em Ciência da Computação	Santosa, Miyazaki e Han, 2021	Avaliação de estratégias de classificação direta e indireta, utilizando técnicas de aprendizado de máquina e <i>deep learning</i> para expandir ontologias existentes.
Sumarização Extrativa de Múltiplos Documentos	Jalil, Nasir e Nasir, 2021	Revisão de técnicas de sumarização extrativa, comparando metodologias e destacando futuras linhas de pesquisa.
Geração de Ontologias de Eventos de Redes Sociais	Rajaonarivo, Mine e Arakawa, 2021	Deteção de eventos locais por meio da análise de movimentação de indivíduos e postagens em redes sociais, utilizando cinco fases distintas de processamento e avaliação baseada em métricas do OntoQA.
Metodologia para Mundos Institucionais Interconectados	Sattar <i>et al.</i> , 2021	Uso de fases de construção colaborativa, integração, versionamento e aplicação prática em gerenciamento de resíduos, destacando-se pela internacionalização e localização.
Learn2Construct	Khemiri <i>et al.</i> , 2021	Construção de ontologias baseada em LDA, validada por experimentos comparativos mostrando resultados promissores.
Sistema de Análise de Avaliações Online	Hananto, Serdült e Kryssanov, 2021	Utilização do algoritmo LDA para rotulação automática de documentos no turismo, com precisão de 70% na classificação de avaliações online.
Criação de Ontologias de Atividades com Crowdsourcing e Gamificação	Sadeghianasl <i>et al.</i> , 2021	Envolvimento de especialistas do domínio em uma abordagem <i>gamificada</i> para criar ontologias de alta qualidade, formalizadas em OWL.
Norms2Onto	Drissi <i>et al.</i> , 2021	Incorporação de técnicas de aprendizado de máquina para construir uma ontologia de dados financeiros, comparada à classificação manual e com módulo de visualização.

Fonte: elaborado pelas autoras (2024).



3.1 METODOLOGIAS DE APRENDIZAGEM DE ONTOLOGIAS

Nesta seção, são apresentadas as metodologias identificadas na revisão de literatura. Em 2018, Deb *et al.* propuseram o uso de taxonomias para criação automática de ontologias no sistema agrícola, diferenciando-se dos métodos convencionais que utilizam dados não estruturados. A validação experimental no domínio agrícola mostrou que essa abordagem pode aprimorar significativamente a precisão e a eficiência na gestão de dados agrícolas.

Outro estudo de 2018 foi desenvolvido por Yadav *et al.*, que criaram a aplicação *EasyOnto*, uma plataforma colaborativa que facilita a construção de ontologias por meio de uma interface gráfica intuitiva. A colaboração entre usuários permite a criação e o aprimoramento contínuo das ontologias, promovendo uma abordagem democrática e inclusiva.

Também em 2018, Enea *et al.* apresentaram uma metodologia baseada na cognição humana para gerenciar incertezas na aprendizagem de ontologias, utilizando extração automática de triplas e validação humana. Essa abordagem garante a precisão e a validade das ontologias, integrando o raciocínio humano no processo automatizado.

Ainda em 2018, Rodriguez e Aguilar propuseram um sistema para extração de conhecimento de textos HTML não estruturados, organizando-os semanticamente em um Grafo de Aprendizagem. Essa metodologia é crucial para transformar dados não estruturados em conhecimento organizado e acessível.

Malchiodi, Da Costa Pereira e Tettamanzi (2018) aplicaram modelos de regressão para prever pontuações de axiomas em ontologias, mostrando resultados satisfatórios. Essa abordagem é essencial para a validação de ontologias complexas, garantindo sua precisão e confiabilidade.

Gyrard, Zimmermann e Sheth, em 2018, propuseram uma metodologia para enriquecer catálogos de ontologias em IoT e cidades inteligentes. A análise e a melhoria de catálogos existentes demonstram a importância de ontologias bem



estruturadas para o desenvolvimento de cidades inteligentes baseadas em tecnologias IoT.

Ainda em 2018, Aggoune desenvolveu uma abordagem semântica para a aprendizagem automática de ontologias a partir de bancos de dados relacionais heterogêneos no domínio de riscos alimentares. Utilizando uma sequência de algoritmos, essa metodologia assegura a extração e a integração precisas de classes e relações, essenciais para a segurança alimentar.

Passando para 2019, Nguyen e Tettamanzi apresentaram uma metodologia para descobrir axiomas a partir de dados RDF, utilizando um algoritmo computacional para identificar axiomas de classes disjuntos. Esse método é essencial para garantir a consistência lógica das ontologias, permitindo a detecção de inconsistências e a inferência de informações implícitas, crucial para a Web Semântica.

Também em 2019, Noura *et al.* introduziram a metodologia KE4WoT para extração de conhecimento, utilizando um *corpus* de 4.500 artigos científicos para identificar tópicos importantes. Essa abordagem não só automatiza a identificação de tópicos, mas também aprimora a interoperabilidade e padronização, essenciais para o desenvolvimento de aplicações semânticas.

Vedula, Maneriker e Parthasarathy, ainda em 2019, apresentaram o *framework* BOLT-K, que utiliza redes neurais para aprender hierarquias ontológicas em novos domínios, transferindo conhecimento existente. Essa metodologia reduz a necessidade de especialistas, tornando o processo de construção de ontologias mais eficiente.

Em 2020, Gómez-Suta, Echeverry-Correa e Soto-Mejía propuseram um método para aprender ontologias a partir de textos sobre o conflito armado colombiano, utilizando agrupamento semântico para extrair e validar conceitos. Essa abordagem é vital para organizar e interpretar informações em contextos específicos e complexos.

Salatino *et al.* (2020) apresentaram uma metodologia para produzir ontologias de tópicos de pesquisa a partir de metadados acadêmicos, oferecendo um portal web para explorar e fornecer *feedback* sobre a ontologia desenvolvida. A capacidade de



focar em ramos específicos ou produzir uma ontologia completa destaca a flexibilidade e utilidade dessa abordagem.

Avançando para 2021, Drissi *et al.* introduziram a metodologia Norms2Onto, que utiliza técnicas de *machine learning* para construção automática e semiautomática de ontologias a partir de um *corpus* textual. Comparada a métodos tradicionais, essa abordagem se destaca por sua eficiência e precisão, automatizando tarefas que tradicionalmente exigiriam intervenção humana intensiva.

Na área de turismo, Hananto, Serdült e Kryssanov, também em 2021, aplicaram a técnica de modelagem de tópicos para automatizar a rotulação de conteúdo, facilitando a construção de ontologias. Essa metodologia não apenas melhora a organização da informação, mas também contribui para sistemas de recomendação mais eficazes, demonstrando a aplicabilidade prática das ontologias em setores específicos.

No domínio da Ciência da Computação, Santosa, Miyazaki e Han (2021) propuseram uma extensão automática de ontologias, avaliando diferentes algoritmos para identificar a sequência mais adequada. Essa pesquisa contribui para a expansão contínua das ontologias, garantindo sua relevância e atualidade.

Sattar *et al.* (2021) destacaram-se por sua flexibilidade para lidar com a complexidade e o tamanho das bases de conhecimento. Essa abordagem é particularmente útil em domínios nos quais a base de conhecimento está em constante expansão e diversificação.

Rajaonarivo, Mine e Arakawa (2021) propuseram uma abordagem dinâmica para a construção de ontologias a partir de publicações no Twitter, utilizando algoritmos que convertem *tweets* em instâncias na ontologia e atualizam dinamicamente com novos dados. Essa metodologia é especialmente relevante para domínios nos quais os dados são gerados continuamente e precisam ser incorporados rapidamente.

Khemiri *et al.* (2021) introduziram uma abordagem para a construção automática de ontologias usando LDA, criando conceitos e inter-relacionamentos a partir de textos.

A estabilidade do modelo, independentemente das mudanças no *corpus*, destaca a robustez dessa metodologia.

Sadeghianasl *et al.* (2021) utilizaram técnicas de *crowdsourcing* e gamificação para aprender ontologias de atividades de processo a partir de *logs* de eventos, promovendo o engajamento dos usuários. Essa metodologia demonstra como a participação ativa dos usuários pode melhorar a qualidade das ontologias.

Ainda em 2021, Jalil, Nasir e Nasir forneceram uma revisão abrangente de técnicas de sumarização extrativa de múltiplos documentos, contribuindo para a gestão de grandes volumes de dados.

Assim, as metodologias revisadas demonstram evolução contínua e diversificada das técnicas para a construção de ontologias, abrangendo uma ampla gama de domínios e abordagens para melhorar a estruturação e interpretação de informações complexas. No Quadro 7, a seguir, são apresentadas as metodologias mais relevantes identificadas, destacando suas contribuições e particularidades:

Quadro 7 – Metodologias identificadas

Abordagem	Referência	Descrição da Metodologia
Modelos de Regressão para Axiomas	Malchiodi, Da Costa Pereira e Tettamanzi, 2018	Previsão de pontuações de axiomas usando técnicas de regressão.
Taxonomias no Sistema Agrícola	Deb <i>et al.</i> , 2018	Uso de taxonomias como entrada para a criação automática de ontologias, diferindo dos métodos convencionais que utilizam dados não estruturados.
Extração de Conhecimento de HTML	Rodriguez e Aguilar, 2018	Transformação de textos não estruturados em HTML em ontologias semânticas.
Cognição Humana e Simulação Conceitual	Enea <i>et al.</i> , 2018	Gestão de incertezas no processo de aprendizagem de ontologias baseada em raciocínio 'e se'.
Ontologias de Bancos de Dados Relacionais	Aggoune, 2018	Aprendizagem automática de ontologias a partir de bancos de dados relacionais heterogêneos.



Construção Colaborativa EasyOnto	–	Yadav <i>et al.</i> , 2018	Aplicação web EasyOnto para construção colaborativa de ontologias.
Enriquecimento de Catálogos de Ontologias	de de	Gyrard, Zimmermann e Sheth, 2018	Enriquecimento de catálogos de ontologias para Internet das Coisas e cidades inteligentes.
Framework BOLT-K		Vedula, Maneriker e Parthasarathy, 2019	Adaptação e transferência de conhecimento de ontologias existentes para novos domínios usando redes neurais.
Algoritmo Computacional para RDF	para	Nguyen e Tettamanzi, 2019	Descoberta automática de axiomas a partir de dados RDF usando um algoritmo computacional.
KE4WoT		Noura <i>et al.</i> , 2019	Extração de conhecimento para a Web of Things usando técnicas de <i>machine learning</i> , analisando 4.500 artigos científicos.
Ontologia de Tópicos de Pesquisa		Salatino <i>et al.</i> , 2020	Produção de ontologias de tópicos de pesquisa a partir de metadados acadêmicos e fontes externas.
Textos em Espanhol		Gómez-Suta, Echeverry-Correa e Soto-Mejía, 2020	Aprendizagem de ontologias a partir de textos sobre o conflito armado colombiano.
Extensão Automática de Ontologia		Santosa, Miyazaki e Han, 2021	Avaliação de diferentes algoritmos para extensão automática de ontologias na ciência da computação.
Extração de Publicações do Twitter	de	Rajaonarivo, Mine e Arakawa, 2021	Construção de uma ontologia específica a partir de publicações em redes sociais, convertendo <i>tweets</i> em instâncias na ontologia.
DSRM		Sattar <i>et al.</i> , 2021	Desenvolvimento de ontologias com base na Design Science Research Methodology (DSRM), adequada para diversos aspectos da construção de ontologias.
Algoritmo LDA		Khemiri <i>et al.</i> , 2021	Construção automática de ontologias a partir de um <i>corpus</i> textual usando o algoritmo LDA.
Modelagem de Tópicos		Hananto, Serdült e Kryssanov, 2021	Utilização da técnica de modelagem de tópicos para automatizar a rotulação de conteúdo no domínio do turismo.



Crowdsourcing Gamificação	e	Sadeghianasl <i>et al.</i> , 2021	Aprendizagem de ontologias de atividades de processo usando <i>logs</i> de eventos com apoio de <i>crowdsourcing</i> e <i>gamificação</i> .
Norms2Onto		Drissi <i>et al.</i> , 2021	Construção automática e semiautomática de ontologias utilizando técnicas de <i>machine learning</i> a partir de um <i>corpus</i> textual.

Fonte: elaborado pelas autoras (2024).

3.2 ETAPAS/FASES DAS METODOLOGIAS DE APRENDIZAGEM DE ONTOLOGIAS

Nesta seção, são apresentadas as etapas e as fases das metodologias identificadas na revisão de literatura. Buscou-se captar tanto os estudos que detalham todas as etapas quanto etapas isoladas, para obter direcionamentos para esta pesquisa.

Em 2018, Deb *et al.* produziram uma metodologia para o domínio agrícola que utiliza taxonomias organizadas em uma base de conhecimento. As etapas envolvem categorização dos dados taxonômicos, pré-processamento de textos, construção de uma biblioteca de algoritmos para extração de relacionamentos e mapeamento ontológico, e definição de restrições do modelo com a ajuda de especialistas.

Yadav *et al.* (2018) desenvolveram a aplicação EasyOnto, uma plataforma colaborativa que facilita a construção de ontologias em sete etapas, desde a escolha do domínio até a criação de vínculos entre categorias. A interface gráfica intuitiva permite que usuários contribuam com novas instâncias, categorias e relações, promovendo uma abordagem inclusiva.

Outro estudo de 2018, de Enea *et al.*, introduziu um método baseado em jogos interrogativos para validar ontologias, utilizando raciocínio indutivo e dedutivo. A interação entre a aplicação e o operador humano melhora o processo de aprendizagem de ontologias.



Rodriguez e Aguilar (2018) desenvolveram uma aplicação para extração de conhecimento de textos não estruturados, utilizando técnicas de processamento de linguagem natural para separar textos em frases, analisar entidades e relações, e atualizar um Grafo de Aprendizagem.

Gyrard, Zimmermann e Sheth (2018) propuseram uma metodologia para enriquecer catálogos de ontologias em IoT e cidades inteligentes, dividida em cinco etapas: investigação de novos domínios, atualização do dicionário, atualização do catálogo RDF, atualização da página HTML e desenvolvimento de aplicações baseadas no catálogo.

Ainda em 2018, Aggoune desenvolveu uma abordagem semântica para a aprendizagem automática de ontologias a partir de bancos de dados relacionais heterogêneos no domínio de riscos alimentares. Utilizando uma sequência de algoritmos, essa metodologia assegura a extração e a integração precisas de classes e relações, essenciais para a segurança alimentar.

Nguyen e Tettamanzi (2019) apresentaram uma metodologia para descobrir axiomas a partir de dados RDF, utilizando um algoritmo computacional para identificar axiomas de classes disjuntos. Esse método é essencial para garantir a consistência lógica das ontologias, permitindo a detecção de inconsistências e a inferência de informações implícitas, crucial para a Web Semântica.

No domínio da Web of Things, Noura *et al.* (2019) detalharam a metodologia KE4WoT em dez etapas, desde a seleção de ontologias até a identificação de tópicos populares, demonstrando a aplicabilidade prática em diversos domínios.

Ainda em 2019, Vedula, Maneriker e Parthasarathy (2019) descreveram o *framework* BOLT-K, que identifica pares de conceitos e as relações entre eles, usando dados anotados para construir ontologias completas. A abordagem é promissora, mas necessita de um detalhamento mais exaustivo das etapas.

No ano de 2020, foram recuperados dois estudos. O primeiro, de Gómez-Suta, Echeverry-Correa e Soto-Mejía (2020), no qual os autores propuseram um método



para aprendizagem de ontologias no contexto do conflito armado colombiano, utilizando técnicas de agrupamento semântico para extrair e validar conceitos. A proposta é acessível e aplicável a dados e domínios com recursos de conhecimento estruturados limitados. O segundo, de Salatino *et al.* (2020), que detalharam uma metodologia para gerar ontologias de tópicos de pesquisa, processando metadados acadêmicos e fontes externas. A criação de um portal para explorar e fornecer *feedback* sobre a ontologia desenvolvida demonstra um compromisso com a evolução contínua e a participação da comunidade científica.

Mas foi em 2021 que se obteve um maior número de estudos relacionados à metodologia de aprendizagem de ontologias, somando-se oito documentos no total. Drissi *et al.* (2021) introduziram a metodologia Norms2Onto, que utiliza técnicas de *machine learning* para construção automática e semiautomática de ontologias a partir de um *corpus* textual. Comparada a métodos tradicionais, essa abordagem se destaca por sua eficiência e precisão, automatizando tarefas que tradicionalmente exigiriam intervenção humana intensiva.

Os autores Hananto, Serdült e Kryssanov (2021) propuseram um padrão de modelagem de tópicos no setor de turismo, dividido em três partes: descoberta de tópicos, etiquetagem automatizada e treinamento de classificadores. Os tópicos são mapeados em um modelo de conhecimento construído a partir de uma revisão de literatura e uma ontologia geral existente, utilizando uma taxonomia da DBPedia para identificar classes de destinos turísticos.

Utilizando a *Design Science Research*, Sattar *et al.* (2021) detalharam uma abordagem *Methodology* (DSRM), dividida em três estágios: pré-conceitualização, conceitualização e pós-conceitualização. Cada estágio abrange várias fases, como estruturação de um Gráfico de Gantt, definição de ferramentas, aquisição de conhecimento, validação, construção colaborativa e manutenção.

Rajaonarivo, Mine e Arakawa (2021) apresentaram uma metodologia de seis fases para extração de ontologias a partir de dados do Twitter, incluindo coleta de *tweets*, descoberta de entidades, filtragem de entidades, enriquecimento de informações,

geração/atualização de ontologias e avaliação. A clareza das fases facilita a replicação e a validação da metodologia.

No domínio da Ciência da Computação, Santosa, Miyazaki e Han (2021) discutiram a extensão automática de ontologias, composta por cinco etapas: extração de tópicos, criação de um gráfico de tópicos, estruturação de um dicionário de tópicos, análise do nível dos tópicos e filtragem de *links* inadequados.

Khemiri *et al.* (2021) propuseram uma abordagem para construção automática de ontologias usando LDA, dividida em cinco módulos: pré-processamento, extração de termos, modelagem de tópicos, extração de conceitos e relações, e visualização de ontologias. A metodologia destaca-se pela robustez e capacidade de responder a questões básicas sobre a ontologia.

Já Sadeghianasl *et al.* (2021) utilizaram técnicas de *crowdsourcing* e gamificação para aprender ontologias de atividades de processo a partir de *logs* de eventos, promovendo o engajamento dos usuários. Essa metodologia demonstra como a participação ativa dos usuários pode melhorar a qualidade das ontologias.

Por fim, Jalil, Nasir e Nasir (2021) forneceram uma revisão abrangente de técnicas de sumarização extrativa de múltiplos documentos, contribuindo para a gestão de grandes volumes de dados. A seguir, no Quadro 8, é possível identificar um consolidado das etapas e fases identificadas nas metodologias referenciadas.

Quadro 8 – Etapas e fases das metodologias

Abordagem	Referência	Detalhamento das Etapas/Fases
Uso de Taxonomias no Sistema Agrícola	Deb <i>et al.</i> , 2018	1. Categorização dos dados taxonômicos. 2. Pré-processamento de textos. 3. Extração de relacionamentos e mapeamento da ontologia. 4. Construção de classes e subclasses. 5. Identificação das propriedades. 6. Indicação das restrições do modelo.



Extração de Conhecimento de HTML	Rodriguez e Aguilar, 2018	1. Separação do texto em frases. 2. Análise de frases (reconhecimento de nomes próprios, análise morfosintática, reconhecimento de entidades e relacionamentos). 3. Atualização do Grafo de Aprendizagem.
Extrações e Validação de Informações	Enea <i>et al.</i> , 2018	1. Extração automática de triplas de fontes heterogêneas. 2. Validação por operador humano por meio de representações gráficas e consultas SPARQL.
Geração de Automática de Ontologias de Riscos Alimentares	Aggoune, 2018	1. Geração de classes. 2. Extração de propriedades de tipo de dados. 3. Geração de propriedades de objetos. 4. Integração dos registros. 5. Detecção de modificações nas bases de dados e atualização da ontologia.
Enriquecimento de Catálogos de Ontologias	Gyrard, Zimmermann e Sheth, 2018	1. Investigação de um novo domínio de aplicativo IoT. 2. Atualização do dicionário. 3. Atualização do conjunto de dados do catálogo RDF. 4. Atualização do Catálogo de Ontologia HTML. 5. Desenvolvimento de aplicações baseadas nos conjuntos de dados do catálogo.
Construção – Colaborativa EasyOnto	Yadav <i>et al.</i> , 2018	1. Escolha de domínio. 2. Adição de instâncias, categorias ou relações. 3. Enriquecimento com anotações semânticas. 4. Vinculação de instâncias a categorias. 5. Escolha de relações. 6. Atribuição automática de classes pai. 7. Criação de vínculos entre categorias.
Extração de Axiomas de Classes RDF	Nguyen e Tettamanzi, 2019	1. Projeto da gramática para gerar axiomas. 2. Processo evolutivo: geração aleatória de axiomas candidatos. 3. Identificação de axiomas na forma de expressões lógicas.



KE4WoT	Noura <i>et al.</i> , 2019	1. Seleção de Ontologias. 2. Pré-processamento. 3. Extração de vocabulário. 4. Extração de Termos. 5. Cálculo da Frequência. 6. Algoritmo Word2vec. 7. Agrupamento K-Means. 8. Agregação de resultados. 9. Atribuição do nome do tópico. 10. Identificação do Tópico Popular.
Framework BOLT-K	Vedula, Maneriker e Parthasarathy, 2019	1. Identificação de pares de conceitos relacionados. 2. Filtragem de conceitos. 3. Ampliação das instâncias de treinamento. 4. Aprendizado das relações entre os conceitos-alvo restantes.
Ontologia de de Tópicos de Pesquisa	Salatino <i>et al.</i> , 2020	1. Processamento de informações de metadados acadêmicos e fontes externas. 2. Inferência de relações semânticas entre palavras-chave. 3. Geração automática de ontologias para o domínio da Ciência da Computação.
Aprendizagem de Ontologias no Conflito Armado Colombiano	Gómez-Suta, Echeverry-Correa e Soto-Mejía, 2020	1. Identificação automática de <i>tags</i> de classes gramaticais. 2. Lematização e normalização. 3. Remoção de palavras irrelevantes, pontuação e caracteres numéricos. 4. Padronização de termos semânticos relacionados.
Construção Automática com LDA	Khemiri <i>et al.</i> , 2021	1. Pré-processamento. 2. Extração de termos. 3. Modelagem de tópicos. 4. Extração de conceitos e relações. 5. Visualização de Ontologias.



Norms2Onto	Drissi <i>et al.</i> , 2021	1. Módulo de pré-processamento: limpeza e normalização dos dados. 2. Módulo de aprendizagem: extração de conceitos relevantes. 3. Módulo de visualização: desenvolvimento de interfaces intuitivas.
DSRM para Desenvolvimento de Ontologias	Sattar <i>et al.</i> , 2021	1. Pré-conceitualização: definição de ferramentas, técnicas, recursos, análise de domínio, etc. 2. Conceitualização: definição de endurantes e perdurantes. 3. Pós-conceitualização: construção colaborativa, implementação, verificação, controle de versão, etc.
Modelagem de Tópicos	Hananto, Serdült e Kryssanov, 2021	1. Descoberta de tópicos por meio de análises online. 2. Processo automatizado de etiquetagem. 3. Treinamento de classificadores e avaliação do modelo.
Extração de Publicações do Twitter	Rajaonarivo, Mine e Arakawa, 2021	1. Coleta de <i>tweets</i>. 2. Descoberta de entidades. 3. Filtragem de entidades e <i>tweets</i>. 4. Enriquecimento de entidades. 5. Geração/atualização de ontologias. 6. Avaliação da ontologia.
Extensão Automática de Ontologia	Santosa, Miyazaki e Han, 2021	1. Extração de novos tópicos de resumos. 2. Proposição de um gráfico de tópicos. 3. Estruturação de um dicionário de tópicos. 4. Análise do nível dos novos tópicos. 5. Filtragem de <i>links</i> inadequados e previsão de <i>links</i> ocultos.

Fonte: elaborado pelas autoras (2024).

3.3 DIFERENCIAIS, LACUNAS E PRINCIPAIS RESULTADOS DAS METODOLOGIAS DE APRENDIZAGEM DE ONTOLOGIAS

Nesta seção, são apresentados os diferenciais, as lacunas e os principais resultados das metodologias propostas pelos autores, identificados na pesquisa. Os diferenciais



das metodologias de geração automática de ontologias são evidentes em diversos estudos.

Yadav *et al.* (2018) enfatizam que o sistema proposto pelos autores acelerou a fase de aquisição de ontologias, eliminando a necessidade de especialistas nas fases iniciais. Rodriguez e Aguilar (2018) afirmam que o sistema criado permitiu atualizar ontologias e léxicos simultaneamente. Gyrard, Zimmermann e Sheth (2018) destacam que a metodologia proposta promove a reutilização de ontologias e a interoperabilidade semântica em aplicações de cidades inteligentes. Deb *et al.* (2018) argumentam que sua aplicação pode ajudar na aprendizagem automática de ontologias a partir de textos taxonômicos e pode ser usada em outras áreas devido à sua natureza genérica. Aggoune (2018) destaca a eficiência na geração de ontologias completas, sem erros técnicos, com baixo custo e tempo de aprendizado, além de um protótipo que ocupa menos espaço em disco.

Noura *et al.* (2019) afirmam que o KE4WoT reduz o tempo de desenvolvimento e a carga de trabalho humano, promovendo a reutilização de tópicos e mantendo as ontologias atualizadas. Vedula, Maneriker e Parthasarathy (2019) ressaltam que sua aplicação é flexível e generalizável para aprender ontologias em subdomínios emergentes, reduzindo a necessidade de conhecimento manual e detectando novos tipos de relacionamentos. Nguyen e Tettamanzi (2019) confirmam a precisão e a cobertura de sua abordagem para extração de axiomas, destacando a geração de axiomas complexos.

Gómez-Suta, Echeverry-Correa e Soto-Mejía (2020) afirmam que sua abordagem permite validar ontologias abstratas sem avaliação manual, sendo acessível em termos de tempo e recursos, e aplicável a domínios com poucos recursos de conhecimento estruturado. Salatino *et al.* (2020) detalham uma metodologia para gerar ontologias de tópicos de pesquisa, processando metadados acadêmicos e fontes externas. A criação de um portal para explorar e fornecer *feedback* sobre a ontologia desenvolvida demonstra um compromisso com a evolução contínua e a participação da comunidade científica.



Drissi *et al.* (2021) destacam que a metodologia Norms2Onto extraiu automaticamente 2.000 conceitos de cinco normas contábeis, enquanto a classificação manual extraiu apenas 50. A aplicação também identificou 1.900 conceitos relevantes e 1.860 subcategorias validadas pelo dicionário Investopedia, em contraste com 45 conceitos relevantes identificados manualmente. Hananto, Serdült e Kryssanov (2021) mencionam que a abordagem baseada no modelo LDA alcançou uma precisão média de 70% na rotulagem automatizada, superando a anotação manual. Khemiri *et al.* (2021) destacam que a ontologia desenvolvida facilita a anotação automática de documentos acadêmicos, proporcionando um processo mais consistente. Santosa, Miyazaki e Han (2021) ressaltam que a abordagem por eles proposta foi avaliada empiricamente e por especialistas, demonstrando melhor desempenho global.

Em contrapartida, poucos autores indicaram lacunas em suas metodologias. Hananto, Serdült e Kryssanov (2021) observaram que o modelo LDA não conseguiu dividir revisões em tópicos suficientemente distintos, resultando em anotações incorretas em algumas classes. Vedula, Maneriker e Parthasarathy (2019) apontaram que a aplicação BOLT-K requer intervenção manual para definir domínios relacionados durante o aprendizado de ontologias.

Com relação aos resultados, Drissi *et al.* (2021) mostraram que o desempenho do modelo Norms2Onto superou a classificação manual, apresentando oportunidades para novas pesquisas. Deb *et al.* (2018) argumentam que sua aplicação pode ajudar na aprendizagem automática de ontologias a partir de textos taxonômicos e pode ser usada em outras áreas devido à sua natureza genérica. Nguyen e Tettamanzi (2019) confirmam a precisão e cobertura de sua abordagem para extração de axiomas, destacando a geração de axiomas complexos. Rajaonarivo, Mine e Arakawa (2021) validaram a confiabilidade das informações na ontologia desenvolvida. Santosa, Miyazaki e Han (2021) demonstraram a viabilidade da extensão automática de ontologias no domínio da Ciência da Computação. Vedula, Maneriker e Parthasarathy (2019) ressaltam que sua aplicação é flexível e generalizável para aprender ontologias em subdomínios emergentes, reduzindo a necessidade de conhecimento manual e detectando novos tipos de relacionamentos. Gyrard, Zimmermann e Sheth (2018) destacam que a metodologia proposta promove a reutilização de ontologias e a

interoperabilidade semântica em aplicações de cidades inteligentes. Gómez-Suta, Echeverry-Correa e Soto-Mejía (2020) afirmam que sua abordagem permite validar ontologias abstratas sem avaliação manual, sendo acessível em termos de tempo e recursos, e aplicável a domínios com poucos recursos de conhecimento estruturado. A seguir, no Quadro 9, um resumo dos diferenciais, lacunas e resultados identificados.

Quadro 9 – Diferenciais, lacunas e resultados

Abordagem	Referência	Diferenciais	Lacunas	Resultados
Aprendizado Automático de Ontologias	Aggoune (2018)	Geração de ontologias completas, sem erros técnicos, com baixo custo e tempo de aprendizado.	Não mencionado.	Não mencionado.
EasyOnto	Yadav et al. (2018)	Aceleração da fase de aquisição de ontologias, eliminando a necessidade de especialistas nas fases iniciais.	Não mencionado.	Não mencionado.
Extração de Conhecimento	Rodriguez e Aguilar (2018)	Permissão para atualizar ontologias e léxicos simultaneamente.	Não mencionado.	Não mencionado.
Aprendizagem de Ontologias a partir de Textos Taxonômicos	Deb et al. (2018)	Não mencionado.	Não mencionado.	Aplicação pode ajudar na aprendizagem automática de ontologias a partir de textos taxonômicos.



Reutilização de Ontologias	Gyrard, Zimmermann e Sheth (2018)	Não mencionado.	Não mencionado.	Promoção da reutilização de ontologias e a interoperabilidade de semântica em aplicações de cidades inteligentes.
KE4WoT	Noura et al. (2019)	Redução do tempo de desenvolvimento e da carga de trabalho humano.	Não mencionado.	Não mencionado.
Extração de Axiomas	Nguyen e Tettamanzi (2019)	Não mencionado.	Não mencionado.	Precisão e cobertura de sua abordagem para extração de axiomas, destacando a geração de axiomas complexos.
BOLT-K	Vedula, Maneriker e Parthasarathy (2019)	Não mencionado.	Requisição da intervenção manual para definir domínios relacionados durante o aprendizado de ontologias.	Aplicação flexível e generalizável para aprender ontologias em subdomínios emergentes, detectando novos tipos de relacionamentos.
Validação de Ontologias	Gómez-Suta, Echeverry-Correa e Soto-	Não mencionado.	Não mencionado.	Validação de ontologias abstratas sem



	Mejía (2020)			avaliação manual, sendo acessível em termos de tempo e recursos.
Norms2Onto	Drissi et al. (2021)	Capacidade de extrair automaticamente 2.000 conceitos de cinco normas contábeis.	Não mencionado.	Desempenho do modelo superou a classificação manual, apresentando oportunidades para novas pesquisas.
Modelagem de Tópicos no Turismo	Hananto, Serdült e Kryssanov (2021)	Abordagem baseada no modelo LDA obteve uma precisão média de 70% na rotulagem automatizada.	Modelo LDA não conseguiu dividir revisões em tópicos suficientemente distintos.	Não mencionado.
Extensão Automática de Ontologias	Santosa, Miyazaki e Han (2021)	Avaliada empiricamente e por especialistas, demonstrando melhor desempenho global.	Não mencionado	Viabilidade da extensão automática de ontologias no domínio da Ciência da Computação.
Anotação Automática de Documentos	Khemiri et al. (2021)	Facilitação da anotação automática de documentos acadêmicos, proporcionando um processo mais	Não mencionado	Não mencionado.



		consistente.		
Geração de Ontologias a partir de Tweets	Rajaonarivo, Mine e Arakawa (2021)	Não mencionado	Não mencionado	Confiabilidade das informações na ontologia desenvolvida.

Fonte: elaborado pelas autoras (2024).

4. CONSIDERAÇÕES FINAIS

A análise das pesquisas sobre a aprendizagem de ontologias revela uma carência significativa na descrição detalhada das metodologias adotadas. A maioria dos trabalhos menciona a utilização de técnicas de automatização, sem aprofundar nas etapas específicas. No entanto alguns estudos, como os de Sattar *et al.* (2021) e Noura *et al.* (2019), oferecem descrições mais claras e detalhadas, contribuindo para uma melhor compreensão do processo.

Há uma alta aplicação de algoritmos computacionais, incluindo aprendizagem de máquina, modelagem de tópicos, grafos, modelos de regressão e descoberta de tópicos. Algumas abordagens adotam estratégias de colaboração e *crowdsourcing* para a construção e avaliação de ontologias. Destaca-se que poucas metodologias mencionam o uso de taxonomias ou outros vocabulários controlados, comuns na área de BCI, e raramente se referem ao uso de ontologias de referência ou de topo.

Muitos trabalhos focam na construção de *frameworks* ou aplicações, sem detalhar o passo a passo para automatizar a criação de ontologias. A variabilidade no detalhamento metodológico é ampla: enquanto alguns estudos fornecem descrições detalhadas, outros se concentram apenas em componentes específicos.

A análise comparativa dos estudos revela avanços significativos e áreas para melhorias na aprendizagem de ontologias. Metodologias como Norms2Onto (Drissi *et al.*, 2021) e KE4WoT (Noura *et al.*, 2019) destacam-se pela automatização de processos complexos, reduzindo a necessidade de intervenção manual e garantindo maior precisão e consistência. Porém há limitações, como a dificuldade de distinguir



tópicos distintos (Hananto, Serdült e Kryssanov, 2021) e a necessidade de intervenção manual para definir domínios relacionados (Vedula, Maneriker e Parthasarathy, 2019).

A geração automática de ontologias apresenta potencial significativo, mas desafios como precisão, flexibilidade e aplicabilidade em diversos domínios precisam ser abordados continuamente. Para avançar no campo da aprendizagem de ontologias, é crucial que estudos futuros forneçam descrições mais completas e sistemáticas de suas metodologias, permitindo maior replicabilidade e aplicação prática.

As contribuições deste estudo vão além da proposta de uma metodologia, identificando direções futuras para pesquisas na interseção entre geração automática de ontologias e governo eletrônico. Espera-se que as descobertas inspirem investigações adicionais e estimulem a implementação de estratégias inovadoras no governo eletrônico, aprimorando a eficiência, a transparência e a acessibilidade dos serviços públicos digitais.

AGRADECIMENTOS

A segunda coautora agradece ao Conselho Nacional de Pesquisa e Desenvolvimento; ao Ministério da Educação, pela concessão da bolsa de Produtividade de Pesquisa através do processo número PQ-1D - Processo 313645/2020-5.

A primeira coautora agradece ao Grupo de pesquisa Modelagem Conceitual para Organização e Representação da Informação Hipertextual (MHTX) e ao Serviço Federal de Processamento de Dados (SERPRO).

REFERÊNCIAS

AGGOUNE, Aicha. Automatic ontology learning from heterogeneous relational databases: Application in alimentation risks field. *In*: AMINE, A; MOUHOU, M.; MOHAMED, A.; DJEBBAR, B. (ed.) **Computational Intelligence and Its Applications**. Cham, Alemanha: Springer International Publishing, 2018. p. 199-210.

ALHARBI, Reham; TAMMA, Valentina; GRASSO, Floriana. Characterising the gap between theory and practice of ontology reuse. *In*: PROCEEDINGS OF THE 11TH ON KNOWLEDGE CAPTURE CONFERENCE, 11., 2021, New York, NY, United States. **Proceedings** [...]. New York: Association for Computing Machinery, 2021. p. 217-224.



ARAÚJO, Webert Júnio *et al.* **Proposta metodológica para enriquecimento de ontologias de domínio**. 271 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2021. Disponível em: <<https://repositorio.ufmg.br/handle/1843/38983>>. Acesso em: 12 jul. 2024.

DEB, Chandan Kumar *et al.* A framework for ontology learning from taxonomic data. *In*: AGGARWAL, V. B.; BHATNAGAR, V.; MISHRA, D. K. **Big Data Analytics: Proceedings of CSI 2015**. Singapore: Springer Singapore, 2018. p. 29-37. vol. 654.

DRISSI, Amani *et al.* A New Automatic Ontology Construction Method Based on Machine Learning Techniques: Application on Financial Corpus. *In*: PROCEEDINGS OF THE 13TH INTERNATIONAL CONFERENCE ON MANAGEMENT OF DIGITAL ECOSYSTEMS, 13., 2021, New York, NY, USA. **Proceedings** [...]. New York: Association for Computing Machinery, 2021. p. 57-61.

ENEA, Roberto *et al.* How to Support Human Operator in "Uncertainty" Managing during the Ontology Learning Process. *In*: COMPANION PROCEEDINGS OF THE THE WEB CONFERENCE 2018 (WWW '18), 18., 2018, Lyon, France. **Proceedings** [...] Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2018. p. 1147-1154.

GÓMEZ-SUTA, Manuela; ECHEVERRY-CORREA, Julián D.; SOTO-MEJÍA, José A. Semi-automatic extraction and validation of concepts in ontology learning from texts in Spanish. *In*: PROCEEDINGS OF THE 10TH INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE, MINING AND SEMANTICS. 10., 2020, New York, NY, United States. **Proceedings** [...]. New York: Association for Computing Machinery, 2020. p. 7-16.

GYRARD, Amelie; ZIMMERMANN, Antoine; SHETH, Amit. Building IoT-based applications for smart cities: How can ontology catalogs help?. **IEEE Internet of Things Journal**, United States, v. 5, n. 5, p. 3978-3990, 2018.

GRUBER, Thomas R. 1993. A Translation Approach to Portable Ontology Specifications. **Knowledge Acquisition**, California, v.5, p. 199–220, 1993.

HANANTO, Valentinus R.; SERDÜLT, Uwe; KRYSSANOV, Victor. A Tourism Knowledge Model through Topic Modeling from Online Reviews. *In*: INTERNATIONAL CONFERENCE ON COMPUTING AND DATA ENGINEERING, 7., 2021 Phuket, Thailand, 15–17 January 2021. **Proceedings** [...]. New York: Association for Computing Machinery, 2021. p. 87-93.

JALIL, Zakia; NASIR, Jamal Abdul; NASIR, Muhammad. Extractive Multi-Document Summarization: A Review of Progress in the Last Decade. **IEEE Access**, United States, v. 9, p. 130928-130946, 2021.

KHEMIRI, Ahmed *et al.* Learn2Construct: an automatic ontology construction based on LDA from textual data. *In*: PROCEEDINGS OF THE 13TH INTERNATIONAL CONFERENCE ON MANAGEMENT OF DIGITAL ECOSYSTEMS, 13., 2021, New



York, NY. **Proceedings** [...]. New York: Association for Computing Machinery, 2021. p. 49-56.

KITCHENHAM, Barbara. Procedures for performing systematic reviews. **Keele, UK, Keele University**, United Kingdom, v. 33, n. 2004, p. 1-26, 2004. Disponível em: <<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=29890a936639862f45cb9a987dd599dce9759bf5>>. Acesso em: 12 jul. 2024.

MALCHIODI, Dario; DA COSTA PEREIRA, Célia; TETTAMANZI, Andrea GB. Predicting the possibilistic score of OWL axioms through support vector regression. *In*: SCALABLE UNCERTAINTY MANAGEMENT: 12TH INTERNATIONAL CONFERENCE, SUM 2018, 12., 2018. Milan, Italy. **Proceedings** [...]. Milan, Italy: Springer International Publishing, October 3-5, 2018. p. 380-386.

MALCHIODI, Dario; TETTAMANZI, Andrea GB. Predicting the possibilistic score of OWL axioms through modified support vector clustering. *In*: PROCEEDINGS OF THE 33RD ANNUAL ACM SYMPOSIUM ON APPLIED COMPUTING, 33., 2018. Pau, France. **Proceedings** [...]. New York: Association for Computing Machinery, 2018. p. 1984-1991.

NGUYEN, Thu Huong; TETTAMANZI, Andrea GB. An evolutionary approach to class disjointness axiom discovery. *In*: IEEE/WIC/ACM INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE, 2019, Greece. **Proceedings** [...]. Greece: IEEE, 2019. p. 68-75.

NOURA, Mahda *et al.* Automatic knowledge extraction to build semantic web of things applications. **IEEE Internet of Things Journal**, v. 6, n. 5, p. 8447-8454, 2019.

RAJAONARIVO, Landy; MINE, Tsunenori; ARAKAWA, Yutaka. Automatic generation of event ontology from social network and mobile positioning data. *In*: IEEE/WIC/ACM INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE AND INTELLIGENT AGENT TECHNOLOGY, 2021, New York. **Proceedings** [...]. New York: IEEE, 2021. p. 87-94.

RODRIGUEZ, Taniana; AGUILAR, Jose. Knowledge extraction system from unstructured documents. **IEEE Latin America Transactions**, United States, v. 16, n. 2, p. 639-646, 2018.

SADEGHIANASL, Sareh *et al.* Process activity ontology learning from event logs through gamification. **IEEE Access**, United States, v. 9, p. 165865-165880, 2021.

SALATINO, Angelo A. *et al.* The computer science ontology: A comprehensive automatically-generated taxonomy of research areas. **Data Intelligence**, United States, v. 2, n. 3, p. 379-416, 2020.

SANTOSA, Natasha C.; MIYAZAKI, Jun; HAN, Hyoil. Automating computer science ontology extension with classification techniques. **IEEE Access**, United States, v. 9, p. 161815-161833, 2021.



SATTAR, Abdul *et al.* An improved methodology for collaborative construction of reusable, localized, and shareable ontology. **IEEE Access**, United States, v. 9, p. 17463-17484, 2021.

VEDULA, Nikhita; MANERIKER, Pranav; PARTHASARATHY, Srinivasan. Bolt-k: Bootstrapping ontology learning via transfer of knowledge. *In: THE WORLD WIDE WEB CONFERENCE*, 19., 2019, New York. **Proceedings** [...]. New York: Association for Computing Machinery, 2019. p. 1897-1908.

YADAV, Usha *et al.* EasyOnto: A Collaborative Semiformal Ontology Development Platform. *In: PANIGRAHI, B., HODA, M., SHARMA, V., GOEL, S. (ed.). Nature Inspired Computing: Proceedings of CSI 2015*. Singapore: Springer Singapore, 2018. p. 1-11.

APÊNDICE - NOTA DE RODAPÉ

3. Uma das definições mais amplamente reconhecidas de ontologia é a de Gruber (1993), que a descreve como uma especificação explícita de uma conceituação, abrangendo tudo o que pode ser formalmente representado. Em um domínio específico do conhecimento, uma ontologia facilita a associação entre os nomes de entidades e suas descrições, de forma que sejam compreensíveis tanto para seres humanos quanto para sistemas computacionais, garantindo uma representação estruturada e compartilhável desse conhecimento.

Material recebido: 08 de agosto de 2024.

Material aprovado pelos pares: 30 de agosto de 2024.

Material editado aprovado pelos autores: 01 de novembro de 2024.

¹ Mestrado em Gestão e Organização do Conhecimento; Especialização em Gestão Pública; Especialização em Design Instrucional para EaD Virtual; Especialização em Produção em Mídias Digitais; Especialização em Educação a Distância; Graduação em Pedagogia. ORCID: <https://orcid.org/0000-0002-4063-3238>. Currículo Lattes: <http://lattes.cnpq.br/0612391970550169>.

² Orientadora. Professora Titular no Departamento de Organização e Tratamento da Informação, na ECI/UFMG. Pós-doutora em Estudos sobre metodologias para o processo de indexação de documentos textuais, pelo Departamento de Biblioteconomia y Documentación, Facultad de Humanidades, Comunicación y Documentación, Universidad Carlos III, Madrid, Espanha, e pela Escola de Artes e Comunicação da USP. Doutora em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação PPGCI-ECI/UFMG. Mestre em Science in Library Service - Clark Atlanta University. Bibliotecária pela Escola de Biblioteconomia da UFMG. Bolsista de Produtividade em



Pesquisa do CNPq - (PQ) Nível 1D. Bolsista do Programa Pesquisador Mineiro-PPM XII-2018 da FAPEMIG. Pesquisadora no Programa de Pós-Graduação em Gestão Organização do Conhecimento-PPGGOC/ECI/UFMG. Coordenadora do Grupo de Pesquisa MHTX - Modelagem Conceitual para Organização e Representação da Informação Hipertextual, registrado no CNPq desde 2004. ORCID: <https://orcid.org/0000-0003-0735-3856>. Currículo Lattes: <http://lattes.cnpq.br/3183050056105009>.