

Submetido 30/05/2024. Aprovado 24/04/2025

Avaliação: revisão duplo-anônimo

Adoção da Inteligência Artificial no Schema Matching: Um Levantamento Sistemático do Estado da Arte

ADOPTION OF ARTIFICIAL INTELLIGENCE IN SCHEMA MATCHING: A SYSTEMATIC SURVEY OF THE STATE OF THE AR

ADOPCIÓN DE INTELIGENCIA ARTIFICIAL EN LA COINCIDENCIA DE ESQUEMAS SCHEMA MATICHING: UN ESTUDIO SISTEMÁTICO DEL ESTADO DEL ARTE

Ricardo Henricki Dias Borges

Universidade Federal de Goiás

ricardoborges@ufg.br

Valdemar Vicente Graciano Neto

Universidade Federal de Goiás

valdemarneto@inf.ufg.br

Leonardo Andrade Ribeiro

Universidade Federal de Goiás

laribeiro@inf.ufg.br

Resumo

Com a crescente complexidade da integração de dados em razão do aumento em sua quantidade e diversidade, o *Schema Matching* desempenha um papel fundamental. Nesse cenário desafiador, a Inteligência Artificial (IA) surge como uma solução promissora para aprimorar a eficiência do *Schema Matching*. Este artigo apresenta os resultados de um mapeamento sistemático da literatura, investigando as técnicas e os algoritmos de IA mais utilizados em aplicações de *Schema Matching*. Os insights obtidos oferecem orientação valiosa para pesquisadores e profissionais que buscam aprimorar a integração de dados por meio do *Schema Matching*.

Palavras-chave: *Schema Matching*; inteligência artificial; mapeamento sistemático.

Abstract

The following text presents a synopsis of the abstract. Given the increasing intricacy of data integration, attributable to both the proliferation of data and the diversification of its characteristics, Schema Matching is a critical component of this process. In the context of this challenging scenario, the application of Artificial Intelligence. The advent of Artificial Intelligence (AI) has emerged as a promising solution to enhance the efficiency of Schema Matching. The present article This paper presents the findings of a systematic literature review that investigated Artificial Intelligence (AI) techniques and algorithms. This is the most common usage in Schema Matching applications. The insights obtained provide valuable guidance. This text is intended for researchers and professionals who are seeking to improve data integration through Schema Matching.

Keywords: Schema Matching; Artificial Intelligence; systematic review.

Resumen

Con la creciente complejidad de la integración de datos debido al aumento de su cantidad y diversidad, *Schema Matching* juega un papel clave. En este desafiante escenario, la Inteligencia Artificial (IA) emerge como una solución prometedora para mejorar la eficiencia del *Schema Matching*. Este artículo presenta los resultados de un mapeo sistemático de la literatura, investigando las técnicas y algoritmos de IA más utilizados en aplicaciones de *Schema Matching*. Los conocimientos adquiridos proporcionan una valiosa orientación para los investigadores y profesionales que buscan mejorar la integración de datos a través de *Schema Matching*.

Palabras clave: Coincidencia de Esquemas; inteligencia artificial; mapeo sistemático.

Introdução

A integração de dados é crucial para muitas organizações que precisam combinar dados de diferentes fontes para transformá-los em informações. No entanto, a integração de dados pode ser um desafio, especialmente quando os dados provêm de fontes com esquemas diferentes. Realizada de forma desordenada, a integração de dados pode levar a erros e inconsistências. Nesse contexto, o *Schema Matching* se apresenta como uma abordagem essencial na integração de dados que consiste na identificação de correspondências entre esquemas de diferentes fontes de dados (Rahm; Bernstein, 2001). Esse processo consiste na identificação e no estabelecimento de correspondências entre os elementos de dois ou mais esquemas de dados diferentes. Essa correspondência permite a integração e o compartilhamento de informações entre sistemas ou fontes de dados heterogêneas, podendo ser aplicado de forma automática, o que leva a uma integração manual mais eficiente e precisa (Bilke; Naumann, 2005).

A contribuição da Inteligência Artificial (IA) para aprimorar o *Schema Matching* é cada vez mais proeminente. Por meio de técnicas de IA, é possível automatizar grande parte do processo de correspondência de esquemas (*Schema Matching*), reduzindo a dependência de intervenção manual intensiva. Isso não apenas acelera a integração de dados, mas também aprimora a precisão, identificando correspondências sutis que poderiam passar despercebidas de outra forma. A IA também tem a capacidade de aprender com as decisões de mapeamentos anteriores, refinando continuamente suas estratégias e adaptações à medida que novos dados e desafios surgem.

No entanto, um mapeamento sistemático sobre a aplicação da IA no *Schema Matching* é fundamental para compreender e sintetizar de maneira completa o estado atual do conhecimento nessa convergência. Dado o constante desenvolvimento das técnicas de IA, tal mapeamento favorecerá uma visão global de abordagens, metodologias e tendências específicas desse domínio, auxiliando na identificação de lacunas de pesquisa, pontos fortes e limitações.

A principal contribuição deste artigo é a apresentação dos resultados de uma abordagem sistemática para mapear as técnicas de IA aplicadas ao *Schema Matching*. No total, 644 estudos foram inicialmente identificados pela *string* de busca, dos quais 68 foram incluídos com base em critérios específicos de seleção e *snowballing*. Notavelmente, as subáreas de *Deep learning* (DL) e Natural Language Processing (NLP) emergiram como as mais amplamente utilizadas na literatura investigada. Além disso, outras técnicas não diretamente relacionadas à IA também foram aplicadas em conjunto com as técnicas de IA, evidenciando a interdisciplinaridade das abordagens utilizadas.

Este artigo está organizado da seguinte forma: inicialmente, são detalhados os métodos adotados no planejamento e na execução do protocolo do Mapeamento Sistemático da Literatura (MSL); em seguida, são apresentados os resultados obtidos ao longo da condução do MSL e a análise correspondente; posteriormente, descrevem-se as ameaças à validade do estudo; e, por fim, apresentam-se as considerações finais e as propostas para trabalhos futuros.

Metodologia

Para desenvolver esta pesquisa, foi utilizado o Mapeamento Sistemático da Literatura (MSL) de acordo com o protocolo proposto por Fabbri et al. (2013) e Petersen, Vakkalanka e Kuzniarz (2015). O protocolo é composto de três etapas principais: Planejamento, Condução e Publicação dos Resultados, apresentado na Figura 1.

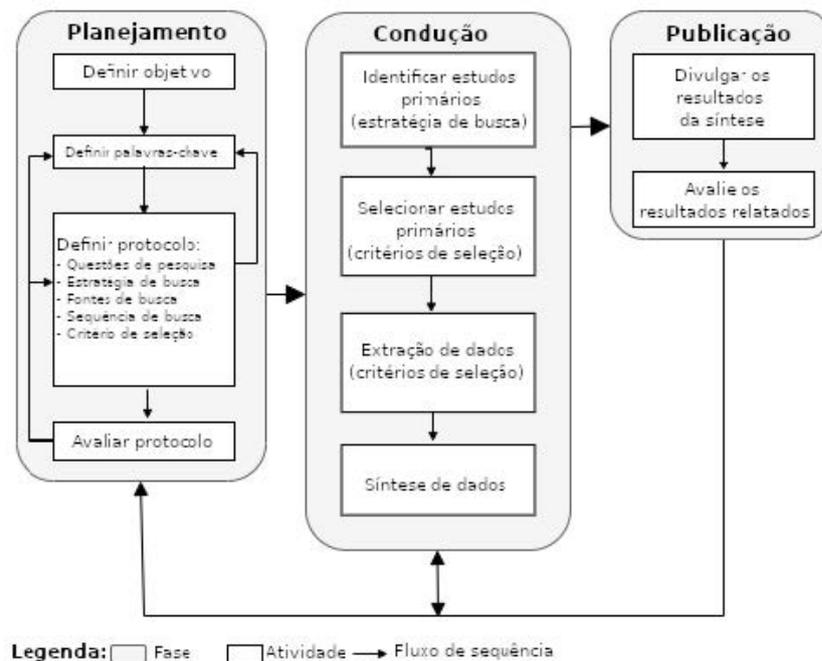


Figura 1 – Fases e atividades do MSL

Fonte: Elaborado pelos(as) autores(as) com base em Fabbri et al. (2013).

Questões de Pesquisa

As questões de pesquisa que expressam os objetivos deste mapeamento foram formuladas seguindo os critérios especificados por PICOC (*Population, Intervention, Comparison, Outcomes, Context*) definido em Budgen e Brereton (2006). Por tratar-se de um mapeamento, apenas PIO foi utilizado. A Tabela 1 exhibe os detalhes.

População	<i>Schema Matching</i> (SM), Inteligência Artificial (IA), <i>Machine Learning</i> (ML), <i>Deep learning</i> (DL), Natural Language Processing (NLP)
Intervenção	Métodos/Técnicas/Tecnologias/Ferramentas/Padrões
Resultados	Técnicas ou algoritmos utilizados

Tabela 1 - Critérios do PIO

Fonte: Elaborado pelo(as) autores(as).

No âmbito da IA, foram exploradas três questões de pesquisa que abordam cada uma das suas subáreas: DL, ML e NLP. Além disso, uma quarta questão de pesquisa foi examinada em relação a técnicas que não estão estritamente vinculadas a nenhuma subárea específica.

A Figura 2 mostra um diagrama de Venn da relação entre as subáreas da IA.

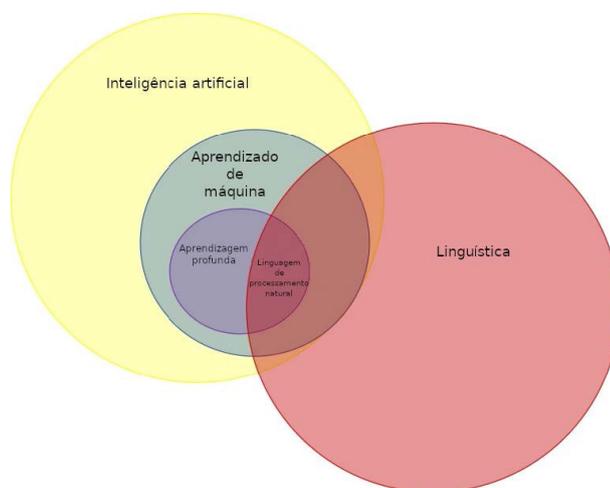


Figura 2 – Diagrama de Venn de subáreas da Inteligência Artificial

Fonte: Elaborado pelo(as) autores(as).

No diagrama da Figura 2 é possível observar que DL está situado dentro do domínio do ML, uma vez que DL é uma abordagem específica que se baseia em redes neurais profundas para o aprendizado de máquina. No entanto, o NLP é representado como uma área separada, pois é um campo especializado que se concentra exclusivamente no processamento e na compreensão da linguagem natural. Embora o NLP utilize técnicas de IA para seu funcionamento, é importante reconhecer sua singularidade devido ao foco específico na linguagem e em desafios linguísticos complexos. A seguir estão as questões de pesquisa (QP) levantadas durante esta fase e sua justificativa.

QP1: Quais técnicas específicas de *Deep learning* são utilizadas na aplicação de *Schema Matching*?

- Procura investigar quais técnicas exclusivamente de *Deep learning* estão sendo aplicadas em *Schema Matching*.

QP2: Quais técnicas específicas¹ de ML têm sido utilizadas na aplicação de *Schema Matching*?

¹ Existem técnicas de ML que não são de DL

- Procura investigar quais técnicas exclusivamente de ML estão sendo aplicadas em *Schema Matching*.

QP3: Quais técnicas específicas² de NLP têm sido utilizadas na aplicação de *Schema Matching*?

- Procura investigar quais técnicas exclusivamente de NLP estão sendo aplicadas em *Schema Matching*.

QP4: Quais técnicas de IA têm sido utilizadas na aplicação de *Schema Matching*?

- Procura investigar quais técnicas exclusivamente de IA que não se enquadram em DL, ML e NLP estão sendo aplicadas em *Schema Matching*.

Identificação dos Estudos

O mapeamento foi conduzido utilizando uma estratégia de busca automática nas bases Scopus, IEEEExplore, ACM e Engineering Village. Além disso, foi aplicada a técnica do *Snowballing* para os estudos selecionados.

STRING DE BUSCA

Fase fundamental para incluir os termos que são pertinentes às questões de pesquisa, ou seja, aqueles relacionados às palavras-chave: inteligência artificial, *deep learning*, *Machine Learning* e *Schema Matching*. Além disso, foi incluído de sinônimos para ampliar a abrangência dos resultados.

A *string* de busca definida para este trabalho foi:

("deep learning") OR ("Machine Learning") OR ("Artificial Intelligence")) AND (("Schema Matching") OR ("ontology matching") OR ("Ontology Alignment"))

Durante a aplicação da String de busca, houve aprimoramento contínuo, realizando ajustes e refinamentos com base nos resultados obtidos e na análise dos estudos encontrados. Ao longo desse processo de calibração da String, foi possível observar a evolução da String de busca, à medida que se faz o refinamento dos termos-chave, considerando sinônimos, variações linguísticas e acrônimos relevantes para a área de estudo.

Como mostrado a seguir:

Versão 1: ("deep learning") AND (("Schema Matching") OR ("ontology matching") OR ("entity matching"));

Versão 2: ("deep learning") AND (("Schema Matching") OR ("ontology matching"));

Versão 3: (("deep learning") OR ("Deep Neural Network")) AND (("schema matching") OR ("ontology matching"));

² Existem técnicas de NLP que não são de ML e DL

Versão 4: ((“*deep learning*”) OR (“*Machine Learning*”) OR (“*Artificial Intelligence*”)) AND ((“*Schema Matching*”) OR (“*ontology matching*”) OR (“*Ontology Alignment*”)).

Critérios de Seleção

A etapa de definição dos critérios de seleção incluem os critérios de inclusão e exclusão, que são estabelecidos para orientar a seleção dos estudos que serão lidos na íntegra. A seguir, são apresentados os critérios de inclusão (**CI**) e os critérios de exclusão (**CE**) utilizados:

CI1: O estudo apresenta algoritmos ou técnicas de *Deep learning* aplicado a *Schema Matching/Ontology matching/Ontology Alignment*;

CI2: O estudo apresenta algoritmos ou técnicas de *Machine Learning* aplicado a *Schema Matching/Ontology matching/Ontology Alignment*;

CI3: O estudo apresenta algoritmos ou técnicas de NLP aplicado a *Schema Matching/Ontology matching/Ontology Alignment*;

CI4: O estudo apresenta conjunto de algoritmos ou técnicas IA aplicado a *Schema Matching/Ontology matching/Ontology Alignment*.

CE1: O estudo não é um estudo primário;

CE2: O estudo não está disponível para acesso gratuito;

CE3: O estudo não está escrito em Inglês ou Português;

CE4: O estudo não foi publicado nos últimos 5 anos;

CE5: O estudo não apresenta algoritmos ou técnicas de Inteligencia Artificial aplicado a *Schema Matching/Ontology matching/Ontology Alignment*;

CE6: O estudo não é um artigo, ou seja, não foi revisado por partes, sendo uma literatura cinza.

Estudos que atenderam pelo menos um critério de inclusão foram incluídos na seleção inicial e os estudos que atendem a pelo menos um critério de exclusão são excluídos da seleção inicial. Inicialmente, foi aplicado o critério de exclusão de limitar os estudos aos últimos 10 anos. No entanto, percebeu-se que o tema em estudo tem passado por uma rápida evolução, especialmente a partir dos anos de 2018 e 2019, com o surgimento de estudos que aplicam diversas técnicas inovadoras. Diante dessa percepção, foi reduzido o escopo do mapeamento para os últimos 5 anos.

Por meio da extração inicial de 644 estudos e importado na ferramenta³, foram eliminados os duplicados (151 estudos). Em seguida, foi realizada a leitura de todos os títulos, resumos e até conclusões de cada estudo, a fim de identificar os estudos que realmente atenderam ao tema proposto. Após isso, foi extraída a seleção inicial dos estudos para realizar a leitura completa. Nesta seleção foram extraídos inicialmente

³ A ferramenta utilizada foi a Parsif.al (<https://parsif.al>).

82 estudos. Porém, com a leitura completa percebeu-se que ainda havia estudos que fugiam do tema proposto. Desta forma, os estudos não relevantes foram excluídos, restando ao final 59 estudos para extração.

Extração dos Dados

Durante essa etapa, interrompeu-se o uso da ferramenta, alterando o foco da manipulação dessa extração de dados através de planilhas, com o propósito de obter uma perspectiva sobre a evolução das seleções ao longo do protocolo. Essas planilhas foram criadas a partir da plataforma Parsif.al. Durante o processo de extração, todas as questões de pesquisa foram abordadas, e os estudos foram categorizados com base nas respostas para cada uma delas.

2.5. *Snowballing*

Com os estudos identificados, foi aplicado o método *Snowballing*, uma técnica de revisão sistemática que consiste em explorar tanto as referências de estudos selecionados (busca retroativa) quanto os trabalhos que os citaram (busca prospectiva). Esse processo permitiu analisar as referências bibliográficas dos estudos já encontrados e identificar trabalhos que os mencionavam, resultando na descoberta de 26 novos estudos relevantes que inicialmente não haviam sido incluídos. Esses 26 estudos passaram pelo mesmo protocolo de análise, e, ao final, apenas nove foram considerados aptos para o trabalho em questão, totalizando 68 estudos. A aplicação do método de *Snowballing* permitiu não apenas ampliar a base de estudos, mas também garantir uma cobertura mais completa e identificar novas fontes relevantes que enriqueceram a pesquisa. Além disso, a técnica facilitou a descoberta de conexões entre trabalhos e a compreensão da evolução das pesquisas na área.

Resultados

Os estudos relevantes estão organizados em tabelas durante as próximas subseções, onde é possível visualizar a referência completa do estudo e um código de identificação para os estudos acrescido de um valor numérico, que será utilizado como referência.

Quais técnicas ou algoritmos de *Deep learning* têm sido utilizados na aplicação de *Schema Matching*?

Entre as técnicas ou algoritmos da *Deep learning* encontradas destacam-se as Redes Neurais Siameses (SNN) com 23,3%, Redes de Memória de Curto Prazo Longa (LSTM) com 26,7% e Redes Neurais Convolucionais (CNN) com 6,7%. Além disso, o Multi-layer Perceptron (MLP) com 13,3%, Rede Neural Recorrente (RNN) e Gráfico de Redes Convolucionais (GCN) com 10,0%. Outros como a Rede Neural Recursiva (RvNNs), Competitive Learning e Multi-Input com 3,3%. Veja na Figura 3.

As proporções indicam que as LSTM e as SNNs são as técnicas de *Deep learning* mais prevalentes no contexto de *Schema Matching*, seguidas por MLPs, RNNs, GCNs e CNNs. Essas informações sugerem quais técnicas são mais frequentemente empregadas ou consideradas úteis para o *Schema Matching*.

SNN	CNN	LSTM	MLP	RNN	GCN	RvNNs, Competitive Learning e MIMO
E01, E02, E03, E04, E05, E06 e E07	E08 e E09	E10, E11, E12, E13, E14, E15 e E16	E17, E18, E19 e E08	E12, E13 e E16	E09, E20 e E21	E14, E07 e E22

Tabela 2 - Estudos selecionados da questão: Quais técnicas ou algoritmos de *Deep learning* têm sido utilizados na aplicação de *Schema Matching*?

Fonte: Elaborado pelos(as) autores(as).



Figura 3 – Técnicas de *Deep learning* que têm sido utilizados no *Schema Matching*

Fonte: Elaborado pelos(as) autores(as).

Identificador	Referência
E01	(Iyer; Agarwal; Kumar, 2020b)
E02	(Iyer; Agarwal; KumaR, 2021)
E03	(Srinivas; Gale; Dolby, 2018)
E04	(Sun; Takeuchi; Yamasaki, 2020)
E05	(Chen <i>et al.</i> , 2021)
E06	(Iyer; Agarwal; Kumar, 2020a)
E07	(Xue <i>et al.</i> , 2021b)
E08	(Bento; Zouaq; Gagnon, 2020)
E09	(Wang <i>et al.</i> , 2022)
E10	(Jiang; XUE, 2020)
E11	(Maji; Rout; Choudhary, 2021)
E12	(Sun; Shen, 2022)
E13	(Mohamed <i>et al.</i> , 2022)
E14	(Chakraborty <i>et al.</i> , 2021)
E15	(Shraga; GAL, 2022)
E16	(Koutras <i>et al.</i> , 2020)
E17	(Xue <i>et al.</i> , 2021a)
E18	(Khan; Gubanov, 2020)
E19)	(Shraga; Gal; Roitman, 2020)
E20	(Hao <i>et al.</i> , 2021)
E21	(Jurisch; Iglar, 2019)
E22	(Hulsebos <i>et al.</i> , 2019)

Tabela 3 - Referência dos estudos selecionados para a questão: Quais técnicas ou algoritmos de *Deep learning* têm sido utilizados na aplicação de *Schema Matching*?

Fonte: Elaborado pelos(as) autores(as).

Quais técnicas ou algoritmos de *Machine Learning* têm sido utilizados na aplicação de *Schema Matching*?

A Figura 4 apresenta uma variedade de técnicas e algoritmos de ML identificados. Entre as técnicas mais comuns, destacam-se o classificador Naive Bayes, com 20,7%, a árvore de decisão (Decision Tree) com 17,2% e a floresta aleatória (Random Forest), com 13,8%. Além disso, como k-means com 13,8%, JRip, SVM com 3,4%, 6,9% respectivamente. Regressão com 10,3%, Adaboost com 6,9% e K-Nearest com 3,4%. Os resultados desses estudos também evidenciam a eficácia do uso de técnicas e algoritmos de ML em aplicações de *Schema Matching*, o que pode ajudar a melhorar a qualidade e a precisão de sistemas que lidam com dados heterogêneos.

Naive Bayes	Decision Tree	Random Forest	k-means	JRip, SVM e C4.5	Adaboost	K-Nearest	Regression, Gaussian mixture e Algoritmo evolutivo
E23, E24, E25, E26, E27, E28 e E29	E23, E30, E31 e E32	E33, E23, E34 e E35	E36, E28, E37 e E38	E26 e E31	E31 e E23	E23 e E31	E33, E23, E24, E67 e E68

Tabela 4 - Estudos selecionados da questão: Quais técnicas ou algoritmos de *Machine Learning* têm sido utilizados na aplicação de *Schema Matching*?
 Fonte: Elaborado pelos(as) autores(as).

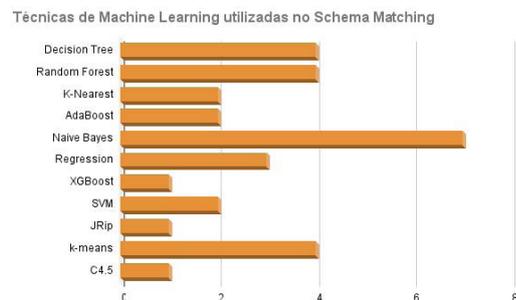


Figura 4 – Técnicas de *Machine Learning* utilizados no *Schema Matching*
 Fonte: Elaborado pelos(as) autores(as).

Identificador	Referência
E23	(Lima <i>et al.</i> , 2020)
E24	(Bulygin, 2018)
E25	(Xue; Chen; Liu, 2021)
E26	(Laadhar <i>et al.</i> , 2019a)
E27	(Schmidts <i>et al.</i> , 2019)
E28	(Berlin; Motro, 2002)
E29	(Nikovski <i>et al.</i> , 2012)
E30	(Amrouch; Mostefai; Fahad, 2016)
E31	(Nezhadi; Shadgar; Osareh, 2011)
E32	(Rodrigues <i>et al.</i> , 2015)

E33	(Bulygin; Stupnikov, 2019)
E34	(Nkisi-Orji <i>et al.</i> , 2019)
E35	(Rodrigues; Silva, 2021)
E36	(Jiménez-Ruiz <i>et al.</i> , 2018b)
E37	(Belhadi <i>et al.</i> , 2023)
E38	(Li; Liu; Zhang, 2005)
E67	(Xue; Chen; Ren, 2019)
E68	(Przyborowski <i>et al.</i> , 2021)

Tabela 5 - Referência dos estudos selecionados para a questão: Quais técnicas ou algoritmos de *Machine Learning* têm sido utilizados na aplicação de *Schema Matching*?

Fonte: Elaborado pelos(as) autores(as).

Quais técnicas ou algoritmos de NLP têm sido utilizados na aplicação de *Schema Matching*?

A Figura 5 relata diversas técnicas e algoritmos de Processamento de Linguagem Natural (NLP). Entre as técnicas mais comuns encontradas destaca-se o BERT (Bidirectional Encoder Representations from Transformers), um modelo de linguagem pré-treinado baseado na arquitetura Transformers, com 17,6% dos estudos. Além do BERT, outros algoritmos de NLP também foram identificados. Por exemplo, o Word2Vec com 41,2%, enquanto o FastText com 11,8%, o GloVe com 11,8%, TransE e StarSpace com 5,9%. Byte-Pair Encoding (BPE) e abordando também o uso de análise de sentimentos, com 2,9% e também técnicas de embedding sem especificar o algoritmo.

Portanto, o Word2Vec é a técnica mais amplamente utilizada, seguida pelo BERT, FastText e GloVe, com outras técnicas também sendo relevantes, embora com menor frequência de uso.

BERT	Word2Vec	FastText	GloVe	TransE	StarSpace	BPE	Embeddings (Sem especificar)
E11, E39, E40, E41, E42, E43	E44, E45, E46, E47, E34, E48, E42, E49, E50, E21, E51, E16, E52 e E53	E40, E54, E42 e E51	E55, E42, E51 e E52	E21, E56	E36 e E57	E58	E01, E02, E06 e E59

Tabela 6 - Estudos selecionados da questão: Quais técnicas ou algoritmos de NLP têm sido utilizados na aplicação de *Schema Matching*?

Fonte: Elaborado pelos(as) autores(as).

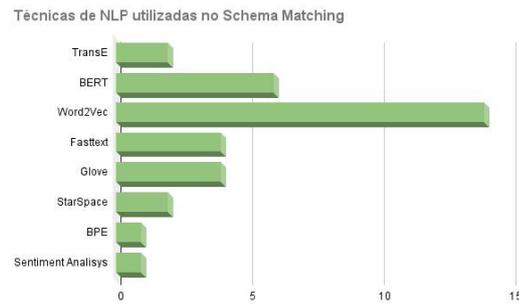


Figura 5 – Técnicas de NLP utilizados no *Schema Matching*

Fonte: Elaborado pelos(as) autores(as).

Identificador	Referência
E44	(Hertling; Portisch; Paulheim, 2020)
E45	(Bulygin, 2018)
E46)	(Teslya; Savosin, 2019)
E47	(Li, 2020a)
E34	(Nkisi-Orji <i>et al.</i> , 2019)
E48	(Nozaki; Hochin; Nomiya, 2019)
E42	(Pan; Pan; Monti, 2022)
E49	(Li, 2020b)
E50	(Chen <i>et al.</i> , 2021)
E21	(Jurisch; Iglar, 2019)
E51	(Cappuzzo, 2020)
E16	(Koutras <i>et al.</i> , 2020)
E52	(Hättasch <i>et al.</i> , 2022)
E53	(ZhanG <i>et al.</i> , 2014)
E40	(Yorsh <i>et al.</i> , 2022)
E54	(Dhouib; Zucker; Tettamanzi, 2019)
E55	(Ayala <i>et al.</i> , 2022)
E56	(Li <i>et al.</i> , 2019)
E36	(Jiménez-Ruiz <i>et al.</i> , 2018b)
E57	(Jiménez-Ruiz <i>et al.</i> , 2018a)
E58	(Zhang <i>et al.</i> , 2021)
E01	(Iyer; Agarwal; Kumar, 2020b)
E02	(Iyer; Agarwal; Kumar, 2021)
E06	(Iyer; Agarwal; Kumar, 2020a)
E59	(Li <i>et al.</i> , 2021)

Tabela 7 - Referência dos estudos selecionados para a questão: Quais técnicas ou algoritmos de NLP têm sido utilizados na aplicação de *Schema Matching*?

Fonte: Elaborado pelos(as) autores(as).

Quais técnicas ou algoritmos de Inteligência artificial têm sido utilizados na aplicação de *Schema Matching*?

A Figura 6 mostra que também foram encontrados estudos isolados, nos quais utilizaram algoritmos mais generalistas da própria IA. Em alguns desses estudos, não foi especificado qual o algoritmo utilizado, enquanto outros mencionaram apenas a subárea (ML ou DL) sem detalhar o algoritmo específico, com 71,4%, ML, 14,3% DL e o Artificial Bee Colonies com 14,3%.

IA Sem especificar o algoritmo	Deep learning	Grasshopper Algorithm	Bee Colonies
E60, E61, E62, E63 e E64	E51	E65	E66

Tabela 8 - Estudos selecionados

Fonte: Elaborado pelos(as) autores(as).

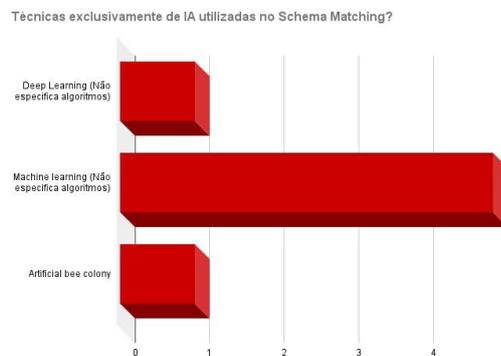


Figura 6 – Técnicas isoladas de IA utilizadas no *Schema Matching*

Fonte: Elaborado pelos(as) autores(as).

Identificador	Referência
E60	(Laadhar <i>et al.</i> , 2019b)
E61)	(Aissaoui; Oughdir, 2020)
E62	(Doan <i>et al.</i> , 2020)
E63	(Mukherjee <i>et al.</i> , 2021)
E64	(Shraga, 2022)
E51	(Cappuzzo, 2020)
E65	(Lv, 2022)
E66	(Rangel <i>et al.</i> , 2015)

Tabela 9 - Quais técnicas ou algoritmos de Inteligência artificial têm sido utilizados na aplicação de *Schema Matching*?

Fonte: Elaborado pelos(as) autores(as).

Discussão

Durante a análise dos estudos, observou-se a importância e o amplo uso de IA aplicada no *Schema Matching* para a integração de dados. A complexidade e o custo de identificar correspondências em uma grande massa de dados pode expandir significativamente o impacto de falhas, tornando o custo e o tempo necessários para resolvê-lo um problema relevante, principalmente quando executada de forma manual.

Além disso, com os dados extraídos do mapeamento, foi constatado a ampla variedade de técnicas de IA aplicadas no *Schema Matching*, sendo uma grande tendência o uso de técnicas de DL e NLP, como na Figura 7. Além disso, também observou-se a presença de abordagens como similaridade e análise de ontologias semânticas junto com técnicas de IA e o uso de mais de uma técnica de IA durante o processo de *Schema Matching*. Essa diversidade de técnicas demonstra que ainda não existe uma única solução ideal para o problema de *Schema Matching*. Pelo contrário, é possível aplicar uma combinação de várias técnicas, adaptando-as ao contexto específico e às necessidades do projeto em questão. A escolha da abordagem mais adequada dependerá das características dos dados, dos requisitos do problema e das metas de qualidade estabelecidas.

Uma das principais lacunas reside na capacidade de compreender semântica e contextos complexos, uma vez que o *Schema Matching* muitas vezes vai além da mera correspondência de palavras. A escalabilidade e o desempenho também são preocupações, especialmente com o aumento na quantidade de dados e esquemas a serem correspondidos. Além disso, manter a correspondência atualizada e automatizada é uma tarefa desafiadora. Isso inclui a necessidade de aprender com dados limitados e medir o desempenho de maneira objetiva. A diversidade na estrutura e no formato dos esquemas, juntamente com a interoperabilidade em ambientes heterogêneos, é outra área que precisa de atenção. Além disso, sistemas que possam aprender com o feedback dos usuários e melhorar continuamente a correspondência de esquemas são necessários.

No entanto, apesar do amplo uso das subáreas de DL e NLP na pesquisa de *Schema Matching*, essas lacunas servem como indicações claras das áreas onde a pesquisa neste campo ainda tem muito a explorar e aprimorar.

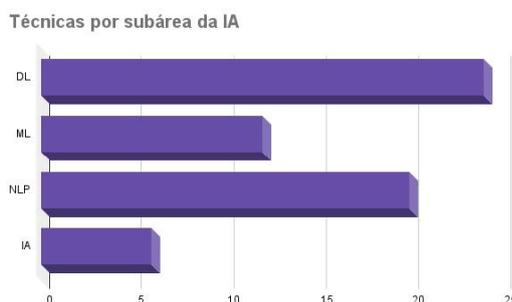


Figura 7 – Técnicas por subárea da IA utilizadas no *Schema Matching*

Fonte: Elaborado pelos(as) autores(as).

Ameaças à validade

Ameaças à validade deste estudo foram identificadas e categorizadas de acordo com Hyman (1982) e Wohlin et al. (2012).

Validade da construção: Pode haver uma possível exclusão de estudos relevantes. Para mitigar esse problema, foram realizadas revisões de cada etapa de condução e extração de dados deste estudo mais de uma vez. A falta da especificidade do tema também foi uma ameaça encontrada. Assim, foram encontrados estudos com foco em técnicas e algoritmos de IA aplicado ao *Schema Matching*, porém sem especificar o algoritmo utilizado. Sendo assim, conseguiu transpor esses resultados identificando pontos aplicados ao *Schema Matching*.

Validade Interna: Podem ter surgido em razão dos métodos de busca escolhidos. Por exemplo, a opção de não incluir algumas bibliotecas digitais pode levar à exclusão de estudos relevantes e ao número relativamente baixo de estudos incluídos. Para mitigar essa ameaça, o protocolo foi previamente avaliado pelos orientadores para identificar possíveis erros.

Validade Externa: Questões externas como a indisponibilidade de estudos foram resolvidas por meio de pesquisa, utilizando o Portal de Periódicos da CAPES. Também utilizou-se a literatura cinza, como Google Acadêmico, Google, Researchgate.

Validade de Conclusão: Possíveis ameaças estão relacionadas ao viés durante a condução e extração de dados, o que pode causar imprecisão na extração de dados, ameaçando a conclusão dos resultados do estudo. Para mitigar essas ameaças, foi apresentado os resultados em uma disciplina de Metodologia Científica e a um grupo de estudo e pesquisa.

Considerações finais

Neste artigo, foram apresentados os resultados de um mapeamento sistemático sobre a aplicação da Inteligência Artificial (IA) no processo de *Schema Matching*. Foram explorados diversos estudos e pesquisas que utilizaram técnicas de IA em DL, ML e NLP, a fim de identificar o uso de técnicas como Redes Neurais Convolucionais (CNN), Multi-layer Perceptron (MLP), dentre outros, para abordar o desafio de *Schema Matching*.

Durante a discussão dos resultados, foi possível observar o amplo uso de IA no contexto do *Schema Matching*. Assim, fica evidente que a IA pode ter um papel crucial na otimização do *Schema Matching*, permitindo a identificação de correspondências entre elementos de esquemas heterogêneos. Essas descobertas enfatizam a necessidade de continuar explorando e aprimorando as técnicas de IA aplicadas ao *Schema Matching*, com o objetivo de compará-las com abordagens não baseadas em IA. Esse enfoque contínuo na evolução das técnicas de IA certamente contribuirá para um avanço significativo no campo do *Schema Matching* e suas aplicações práticas.

Para futuras pesquisas, é importante explorar o desenvolvimento de novos modelos de IA personalizados para abordar os desafios do *Schema Matching*, como o uso de abordagens híbridas que combinam técnicas de IA com métodos tradicionais

de *Schema Matching*, que podem oferecer um potencial significativo. A garantia de segurança e privacidade durante o processo de *Schema Matching* para dados críticos também são desafios adicionais que merecem atenção.

Referências

AISSAOUI, O. E.; OUGHDIR, L. A learning style-based ontology matching to enhance learning resources recommendation. *In: IEEE. 2020 1st international conference on innovative research in applied science, engineering and technology (IRASET)*. [S. l.], p. 1-7, 2020.

AMROUCH, S.; MOSTEFAI, S.; FAHAD, M. Decision trees in automatic ontology matching. *International Journal of Metadata, Semantics and Ontologies*, [s. l.], v. 11, n. 3, p. 180-190, 2016.

AYALA, D.; HERNÁNDEZ, I.; RUIZ, D.; RAHM, E. Leapme: Learning-based property matching with embeddings. *Data & Knowledge Engineering*, Amsterdã, v. 137, 2022.

BELHADI, A.; DJENOURI, Y.; SRIVASTAVA, G.; LIN, J. C.-W. Fast and accurate framework for ontology matching in web of things. *ACM Transactions on Asian and Low-Resource Language Information Processing*, New York, v. 22, n. 5, p. 1-19, 2023.

BENTO, A.; ZOUAQ, A.; GAGNON, M. Ontology matching using convolutional neural networks. *In: SYMPOSIUM ON APPLIED COMPUTING, 35.*, 2020, Brno. *Proceedings [...]*. New York: ACM, 2020.

BERLIN, J.; MOTRO, A. Database schema matching using machine learning with feature selection. *In: SPACCAPIETRA, S.; MARCH, S. T.; KAMBAYASHI, Y. (ed.). Conceptual Modeling: ER 2002*. Berlin; Heidelberg: Springer, 2002.

BILKE, A.; NAUMANN, F. Schema matching using duplicates. *In: INTERNATIONAL CONFERENCE ON DATA ENGINEERING, 21.*, 2005, Tóquio. *Proceedings [...]*. Washington, DC: IEEE Computer Society, 2005. p. 69-80.

BUDGEN, D.; BRERETON, P. Performing systematic literature reviews in software engineering. *In: INTERNATIONAL CONFERENCE ON SOFTWARE ENGINEERING, 28.*, 2006, Xangai. *Proceedings [...]*. New York: ACM, 2006. p. 1051-1052.

BULYGIN, L. Combining lexical and semantic similarity measures with machine learning approach for ontology and schema matching problem. *Procedia Computer Science*, Amsterdã, v. 128, p. 1370-1379, 2018.

BULYGIN, L.; STUPNIKOV, S. A. Applying of machine learning techniques to combine string-based, language-based and structure-based similarity measures for ontology matching. *In: IEEE CONFERENCE OF RUSSIAN YOUNG RESEARCHERS IN ELECTRICAL AND ELECTRONIC ENGINEERING (EIconRus)*, 2019, São Petersburgo. *Proceedings [...]*. Piscataway: IEEE, 2019.

CAPPUZZO, R. *Creating embeddings of heterogeneous relational datasets for data integration tasks*. 2020. Dissertação (Mestrado em Engenharia da Computação) – Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Padova, Pádua, 2020.

CHAKRABORTY, J. *et al.* Onto-connect: Unsupervised ontology alignment with recursive neural network. *In: INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA, 2021, Xi'an. Proceedings [...]*. New York: ACM, 2021. p. 2363-2366.

CHEN, J. *et al.* Augmenting ontology alignment by semantic embedding and distant supervision. *In: HOSE, K. et al. (ed.). The Semantic Web – ISWC 2021*. Cham: Springer, 2021. p. 124-140. (Lecture Notes in Computer Science, v. 12922).

DHOUIB, M. T.; ZUCKER, C. F.; TETTAMANZI, A. G. An ontology alignment approach combining word embedding and the radius measure. *In: HITZLER, P. et al. (ed.). Knowledge Engineering and Knowledge Management*. Cham: Springer, 2019. p. 115-130.

DOAN, A. *et al.* Magellan: toward building ecosystems of entity matching solutions. *Communications of the ACM*, New York, v. 63, n. 8, p. 83-91, ago. 2020.

FABBRI, S. C. P. F. *et al.* Externalising tacit knowledge of the systematic review process. *IET Software*, [s. l.], v. 7, n. 6, p. 298-307, 2013. DOI: 10.1049/iet-sen.2013.0029.

HAO, J. *et al.* Medto: Medical data to ontology matching using hybrid graph neural networks. *In: INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA, 2021, Xi'an. Proceedings [...]*. New York: ACM, 2021. p. 757-770.

HÄTTASCH, B.; TRUONG-NGOC, M.; SCHMIDT, A.; BINNIG, C. *It's ai match: A two-step approach for schema matching using embeddings*. 2022. arXiv:2203.04366.

HERTLING, S.; PORTISCH, J.; PAULHEIM, H. *Supervised ontology and instance matching with melt*. 2020. arXiv:2009.11102.

HULSEBOS, M. *et al.* Sherlock: A deep learning approach to semantic data type detection. *In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY & DATA MINING, 25., 2019, Anchorage. Proceedings [...]*. New York: ACM, 2019. p. 1514-1524.

HYMAN, R. [Resenha de] Quasi-experimentation: Design and analysis issues for field settings. *Journal of Personality Assessment*, v. 46, n. 1, p. 96-97, 1982.

IYER, V.; AGARWAL, A.; KUMAR, H. *Multifaceted context representation using dual attention for ontology alignment*. arXiv:2010.11721, 2020.

IYER, V.; AGARWAL, A.; KUMAR, H. *Multifaceted context representation using dual attention for ontology alignment*.¹ 2020. arXiv:2010.11721. Disponível em: <https://arxiv.org/abs/2010.11721>. Acesso em: 15 set. 2025.

IYER, V.; AGARWAL, A.; KUMAR, H. Veealign: multifaceted context representation using dual attention for ontology alignment.² *In: THE WEB CONFERENCE, 2021, Liubliana. Companion Proceedings [...].* New York: ACM, 2021. p. 317-321.

JIANG, C.; XUE, X. Matching biomedical ontologies with long short-term memory networks. *In: IEEE INTERNATIONAL CONFERENCE ON BIOINFORMATICS AND BIOMEDICINE (BIBM), 2020, Seul. Proceedings [...].* Piscataway: IEEE, 2020. p. 1845-1852.

JIMÉNEZ-RUIZ, E. *et al.* Breaking-down the ontology alignment task with a lexical index and neural embeddings. 2018. arXiv:1805.12402. Disponível em: <https://arxiv.org/abs/1805.12402>. Acesso em: 15 set. 2025.

JIMÉNEZ-RUIZ, E. *et al.* We divide, you conquer: from large-scale ontology alignment to manageable sub-tasks with a lexical index and neural embeddings. *In: INTERNATIONAL WORKSHOP ON ONTOLOGY MATCHING, 13., 2018, Monterey. Proceedings [...].* Aachen: CEUR-WS.org, 2018. p. 13-24. (CEUR Workshop Proceedings, v. 2288).

JURISCH, M.; IGLER, B. Graph-convolution-based classification for ontology alignment change prediction. *In: INTERNATIONAL WORKSHOP ON ONTOLOGY MATCHING, 14., 2019, Auckland. Proceedings [...].* Aachen: CEUR-WS.org, 2019. p. 61-72. (CEUR Workshop Proceedings, v. 2536).

KHAN, R.; GUBANOV, M. Weblens: Towards web-scale data integration, training the models. *In: IEEE INTERNATIONAL CONFERENCE ON DATA ENGINEERING, 36., 2020, Dallas. Proceedings [...].* Piscataway: IEEE, 2020. p. 1757-1760.

KOUTRAS, C. *et al.* Rema: Graph embeddings-based relational schema matching.³ *In: WORKSHOPS OF THE EDBT/ICDT JOINT CONFERENCE, 2020, Copenhagen. Proceedings [...].* Aachen: CEUR-WS.org, 2020. p. 219-226. (CEUR Workshop Proceedings, v. 2592).

LAADHAR, A. *et al.* Partitioning and local matching learning of large biomedical ontologies. *In: INTERNATIONAL CONFERENCE ON EVALUATION OF NOVEL APPROACHES TO SOFTWARE ENGINEERING, 14., 2019, Heraclião. Proceedings [...].* Setúbal: SciTePress, 2019. p. 310-317.

LAADHAR, A. *et al.* Pomap++ results for oaei 2019: fully automated machine learning approach for ontology matching. *In: INTERNATIONAL WORKSHOP ON ONTOLOGY MATCHING, 14., 2019, Auckland. Proceedings [...].* Aachen: CEUR-WS.org, 2019. p. 169-174. (CEUR Workshop Proceedings, v. 2536).

LI, G. Deepfca: Matching biomedical ontologies using formal concept analysis embedding techniques. *In: IEEE INTERNATIONAL CONFERENCE ON BIOINFORMATICS AND BIOMEDICINE (BIBM), 2020, Seul. Proceedings [...].* Piscataway: IEEE, 2020. p. 1829-1836.

- LI, G. Improving biomedical ontology matching using domain-specific word embeddings. *In: IEEE INTERNATIONAL CONFERENCE ON BIOINFORMATICS AND BIOMEDICINE (BIBM), 2020, Seul. Proceedings [...]. Piscataway: IEEE, 2020. p. 1837-1844.*
- LI, W. *et al.* Multi-view embedding for biomedical ontology matching. *In: INTERNATIONAL WORKSHOP ON ONTOLOGY MATCHING, 14., 2019, Auckland. Proceedings [...]. Aachen: CEUR-WS.org, 2019. p. 13-24. (CEUR Workshop Proceedings, v. 2536).*
- LI, X. *et al.* Heterogeneous embeddings for relational data integration tasks. *In: LIN, X.; ZHANG, Y.; ZHANG, W. (ed.). Database Systems for Advanced Applications. Cham: Springer, 2021. p. 45-59. (Lecture Notes in Computer Science, v. 12681).*
- LI, Y.; LIU, D.-B.; ZHANG, W.-M. Schema matching using neural network. *In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND CYBERNETICS, 2005, Guangzhou. Proceedings [...]. Piscataway: IEEE, 2005. p. 4467-4472.*
- LIMA, B. *et al.* Learning reference alignments for ontology matching within and across domains. *In: INTERNATIONAL WORKSHOP ON ONTOLOGY MATCHING, 15., 2020, Atenas. Proceedings [...]. Aachen: CEUR-WS.org, 2020. p. 85-96. (CEUR Workshop Proceedings, v. 2788).*
- LV, Z. An effective approach for large ontology matching using multi-objective grasshopper algorithm. *Scientific Reports, London, v. 12, n. 1, p. 1-17, 2022.*
- MAJI, S.; ROUT, S. S.; CHOUDHARY, S. Dcom: A deep column mapper for semantic data type detection. 2021. arXiv:2106.12871. Disponível em: <https://arxiv.org/abs/2106.12871>. Acesso em: 15 set. 2025.
- MOHAMED, A. *et al.* Schema matching based on deep learning using lstm model. *In: INTERNATIONAL CONFERENCE ON INNOVATIVE RESEARCH IN APPLIED SCIENCE, ENGINEERING AND TECHNOLOGY, 2., 2022, Erode. Proceedings [...]. Piscataway: IEEE, 2022. p. 1-6.*
- MUKHERJEE, D.; BANDYOPADHYAY, A.; CHOWDHURY, R.; BHATTACHARYA, I. Learning knowledge graph for target-driven schema matching. *In: INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA, 2021, Xi'an. Proceedings [...]. New York: ACM, 2021. p. 1351-1364.*
- MUYLAERT, R. Pandemia do novo coronavírus, Parte 6: inteligência artificial (NLP). *Marco Armello, 19 ago. 2020. Disponível em: https://marcoarmello.wordpress.com/2020/08/19/coronavirus6/. Acesso em: 9 jul. 2023.*
- NEZHADI, A. H.; SHADGAR, B.; OSAREH, A. Ontology alignment using machine learning techniques. *International Journal of Computer Science & Information Technology, Chennai, v. 3, n. 2, p. 139-149, 2011.*

- NIKOVSKI, D.; ESEITHER, A.; YE, X.; SHIBA, M.; TAKAYAMA, S. Bayesian networks for matcher composition in automatic schema matching. *In: INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND INFORMATION RETRIEVAL, 2012, Barcelona. Proceedings [...].* Setúbal: SciTePress, 2012. p. 64-73.
- NKISI-ORJI, I. *et al.* Ontology alignment based on word embedding and random forest classification. *In: LAUX, F.; BRUNZEL, M. (ed.). Declarative AI: Theory, Systems, and Applications.* Cham: Springer, 2019. p. 147-163. (Lecture Notes in Computer Science, v. 11867).
- NOZAKI, K.; HOCHIN, T.; NOMIYA, H. Semantic schema matching for string attribute with word vectors. *In: INTERNATIONAL CONFERENCE ON INFORMATION NETWORKING (ICOIN), 2019, Kuala Lumpur. Proceedings [...].* Piscataway: IEEE, 2019. p. 488-493.
- PAN, Z.; PAN, G.; MONTI, A. Semantic-similarity-based schema matching for management of building energy data. *Energies*, Basel, v. 15, n. 23, p. 8894, 2022.
- PETERSEN, K.; VAKKALANKA, S.; KUZNIARZ, L. Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, [s. l.], v. 64, p. 1-18, ago. 2015.
- PRZYBOROWSKI, M.; PABIŚ, M.; JANUSZ, A.; ŚLĘZAK, D. *Schema matching using gaussian mixture models with wasserstein distance.* 2021. arXiv:2111.14244. Disponível em: <https://arxiv.org/abs/2111.14244>. Acesso em: 15 set. 2025.
- RAHM, E.; BERNSTEIN, P. A. A survey of approaches to automatic schema matching. *The VLDB Journal*, [s. l.], v. 10, n. 4, p. 334-350, dez. 2001.
- RANGEL, C. *et al.* An approach for the emerging ontology alignment based on the bees colonies. *In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE (ICCI), 12., 2015, Madri. Proceedings [...].* Atlantis Press, 2015. p. 248-254.
- RODRIGUES, D.; SILVA, A. A study on machine learning techniques for the schema matching network problem. *Journal of the Brazilian Computer Society*, Porto Alegre, v. 27, n. 1, p. 1-32, dez. 2021.
- RODRIGUES, D.; SILVA, A. da; RODRIGUES, R.; SANTOS, E. dos. Using active learning techniques for improving database schema matching methods. *In: IEEE SYMPOSIUM ON COMPUTER-BASED MEDICAL SYSTEMS, 28., 2015, São Carlos. Proceedings [...].* Piscataway: IEEE, 2015. p. 104-109.
- SCHMIDTS, O.; KRAFT, B.; SIEBIGTEROTH, I.; ZÜNDORF, A. Schema matching with frequent changes on semi-structured input files: A machine learning approach on biological product data. *In: INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS, 21., 2019, Heraclião. Proceedings [...].* Setúbal: SciTePress, 2019. v. 1, p. 208-215.

- SHRAGA, R.; GAL, A. Humanal: Calibrating human matching beyond a single task. *In: ACM SIGMOD INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA, 2022, Filadélfia. Proceedings [...].* New York: ACM, 2022. p. 1827-1840.
- SHRAGA, R.; GAL, A. Powarematch: A quality-aware deep learning approach to improve human schema matching. *ACM Journal of Data and Information Quality*, New York, v. 14, n. 3, p. 1-27, set. 2022.
- SHRAGA, R.; GAL, A.; ROITMAN, H. Adnev: Cross-domain schema matching using deep similarity matrix adjustment and evaluation. *Proceedings of the VLDB Endowment*, [s. l.], v. 13, n. 9, p. 1401-1415, 2020.
- SRINIVAS, K.; GALE, A.; DOLBY, J. *Merging datasets through deep learning*. 2018. arXiv:1809.01604. Disponível em: <https://arxiv.org/abs/1809.01604>. Acesso em: 15 set. 2025.
- SUN, C.; SHEN, D. Towards deep entity resolution via soft schema matching. *Neurocomputing*, v. 471, p. 107-117, fev. 2022.
- SUN, J.; TAKEUCHI, S.; YAMASAKI, I. Few-shot ontology alignment model with attribute attentions. *In: IEEE CONGRESS ON EVOLUTIONARY COMPUTATION (CEC), 2020, Glasgow. Proceedings [...].* Piscataway: IEEE, 2020. p. 1-8.
- TESLYA, N.; SAVOSIN, S. Matching ontologies with word2vec-based neural network. *In: ABRAHÃO, S.; ZELENKOV, Y. (ed.). Software and Compilers for Embedded Systems*. Cham: Springer, 2019. p. 115-125. (Lecture Notes in Computer Science, v. 11789).
- WANG, P.; ZOU, S.; LIU, J.; KE, W. Matching biomedical ontologies with gcn-based feature propagation. *Mathematical Biosciences and Engineering*, Springfield, v. 19, n. 8, p. 8479-8504, 2022.
- WOHLIN, C. *et al. Experimentation in software engineering*. Berlin: Springer Science & Business Media, 2012.
- XUE, X.; CHEN, D.; LIU, W. Naive bayesian classifier based semi-supervised learning for matching ontologies. *In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER SUPPORTED COOPERATIVE WORK IN DESIGN (CSCWD), 24., 2021, Dalian. Proceedings [...].* Piscataway: IEEE, 2021. p. 1305-1310.
- XUE, X.; CHEN, J.; REN, A. Interactive ontology matching based on evolutionary algorithm. *In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND SECURITY (CIS), 15., 2019, Macau. Proceedings [...].* Piscataway: IEEE, 2019. p. 1-5.
- XUE, X. *et al.* Artificial neural network based sensor ontology matching technique. *In: IEEE INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND INFORMATION SYSTEMS (ICAIS), 2021, Chongqing. Proceedings [...].* Piscataway: IEEE, 2021. p. 1-6.

XUE, X. *et al.* Matching sensor ontologies through siamese neural networks without using reference alignment. *PeerJ Computer Science*, San Diego, v. 7, e602, 2021.

YORSH, U.; BEHR, A. S.; KOCKMANN, N.; HOLEŇA, M. *Text-to-ontology mapping via natural language processing models*. 2022. arXiv:2209.04944. Disponível em: <https://arxiv.org/abs/2209.04944>. Acesso em: 15 set. 2025.

ZHANG, J.; SHIN, B.; CHOI, J. D.; HO, J. C. *Smat: An attention-based deep learning solution to the automation of schema matching*. In: INTERNATIONAL CONFERENCE ON ASIAN DIGITAL LIBRARIES, 23., 2021, Virtual Event. *Proceedings* [...]. Cham: Springer, 2021. p. 3-17. (Lecture Notes in Computer Science, v. 13127).

ZHANG, Y. *et al.* Ontology matching with word embeddings. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 2014, Pequim. *Proceedings* [...]. Beijing: AAAI Press, 2014. p. 1-6.