The Scientific Temper (2025) Vol. 16 (spl-1): 141-148
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2025.16.spl-1.18
https://scientifictemper.com/
ORIGINAL RESEARCH PAPER

Knowledge graphs for NLP: A comprehensive analysis

Sharayu Mirasdar[1*], Mangesh Bedekar[2]

Abstract

Comprehensive analysis done for this paper examines the blend of knowledge graphs (KGs) and natural language processing (NLP), emphasizing the collective potential of both techniques to improve understanding and processing of textual data amid its rapid growth. KGs provide structured semantic representations that facilitate deeper reasoning and contextual understanding, addressing the limitations inherent in traditional NLP approaches. By consolidating insights from over 79 research papers, the review in-depth explores the definitions, applications, and challenges related to the integration of KGs and NLP, as well as their synergistic applications in multiple domains, such as question answering, sentiment analysis, and text summarization. The review underscores the transformative impact of KGs in bridging unstructured text with structured data, paving the way for innovative methodologies in AI applications. Additionally, it identifies prevailing challenges in the construction and management of KGs while emphasizing the ongoing evolution and promising future of this integrated approach in tackling real-world NLP challenges. The findings aim to benefit both researchers and practitioners in the field, promoting the adoption of KG-based methods across diverse applications.

Keywords - Knowledge graph, Natural language processing, Applications of KGs

Introduction

In recent years, textual data has experienced exponential growth thereby creating a pressing need for the development of advanced solutions to understand, process, and derive meaningful insights from natural language. From conversational agents and search engines to automated summarization and sentiment analysis, natural language processing (NLP) applications have become integral to modern technology. In spite of significant advancements, many NLP systems still face challenges in capturing the deeper semantics, reasoning, and contextual understanding that is considered necessary for human-like comprehension. Knowledge graphs (KGs) have evolved as a popular transformative tool to enhance NLP by providing structured, semantic representations of information. KGs represent entities and their relationships as interconnected graphs, enabling machines to reason and infer beyond surface-level text. The integration of KGs with NLP has shown immense potential across a variety of applications, inclusive of question answering, sentiment analysis, and text summarization. KGs enable machines to go beyond text-based pattern matching and engage in semantic reasoning, answering complex queries and drawing connections between seemingly unrelated pieces of information. The integration of KGs with NLP has opened new avenues for innovation, addressing the limitations of traditional NLP approaches. They have shown immense potential in enhancing NER, question answering systems, text summarization, semantic search, as well as sentiment analysis.

By consolidating insights from over 79 research papers, the literature review offers a thorough understanding of the integration of knowledge graphs and natural language processing. The findings will benefit not only researchers but also practitioners seeking to apply KG-based methods to real-world NLP challenges.

Background and Fundamentals

Knowledge Graphs (KGs)

Graphical representation holds significant importance as it offers a natural and flexible approach to representing complex relationships and interactions between entities.

[1]Assistant Professor, School of Computer Science & Engineering, DES Pune University, Pune, Maharashtra, India.

[2]Professor, School of Computer Science & Engineering, MIT World Peace University, Pune, Maharashtra, India.

***Corresponding Author:** Sharayu Mirasdar, Assistant Professor, School of Computer Science & Engineering, DES Pune University, Pune, Maharashtra, India., E-Mail: sharayu.mirasdar@despu.edu.in

How to cite this article: Mirasdar, S., Bedekar, M. (2025). Knowledge graphs for NLP: A comprehensive analysis. The Scientific Temper, **16**(spl-1):141-148.

Doi: 10.58414/SCIENTIFICTEMPER.2025.16.spl-1.18

Source of support: Nil

Conflict of interest: None.

Received: 19/04/2025 **Accepted:** 17/05/2025 **Published** : 21/05/2025

This quality makes graphs out to be a particularly valuable asset in fields like computer science, mathematics, biology, and social sciences. Their ability to model complex systems, facilitate efficient computation, and power advanced AI applications makes them indispensable in many areas of research and industry.

A Knowledge Graph is specifically a well-structured representation of knowledge in the form of entities and the relationships between them, typically organized as a graph. The use of the term "knowledge graph" can be referenced back to 2012 when Google introduced its Knowledge Graph as part of its search engine to provide richer, contextually aware search results (Singhal A, 2012). After this introduction, a significant increase in the research of Knowledge Graphs can be observed. Since Google's introduction, Knowledge Graphs have also been employed by Facebook, Microsoft Bing, IBM Watson, and eBay (Noy N & Gao Y, 2019). Openly available Knowledge Graphs for DBpedia, YAGO, and Freebase have also emerged since (Paulheim H, 2016). Knowledge Graphs are said to have originated as a new framing attributed to the research on semantic networks, ontologies, and linked data (Hitzler P, 2021).

Defining Knowledge Graphs

During this literature survey, it was observed that there is a lot of discussion about what can be the correct definition of a Knowledge Graph. While a number of definitions that are conflicting at times have emerged, one states the Knowledge Graph definition as *"a graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and whose edges represent potentially different relations between these entities"* (Hogan A & Blomqvist E, 2021). Another definition states that *"Knowledge graphs are large networks of entities, their semantic types, properties, and relationships between entities"* (Kroetsch M, & Weikum G, 2016). They have also been defined as *"a network of all kinds of things which are relevant to a specific domain or to an organization. They are not limited to abstract concepts and relations but can also contain instances of things like documents and datasets"* (Blumauer A, 2014). One definition stated that *"A knowledge graph (i) mainly describes real-world entities and their interrelations, organized in a graph, (ii) defines possible classes and relations of entities in a schema, (iii) allows for potentially interrelating arbitrary entities with each other and (iv) covers various topical domains"*. A definition that describes knowledge graphs as a Resource Description Framework (RDF) graph, states that *"We define a Knowledge Graph as an RDF graph. An RDF graph consists of a set of RDF triples where each RDF triple (s, p, o) is an ordered set of the following RDF terms: a subject $s \in U \cup B$, a predicate $p \in U$, and an object $U \cup B \cup L$. An RDF term is either a URI $u \in U$, a blank node $b \in B$, or a literal $l \in L$" (Färber M, & Bartscherer F & Menne C, & Rettinger A, 2017).* Upon illustrating a knowledge graph architecture, shown in

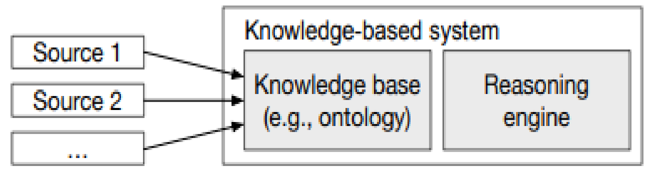

Figure 1: Architecture of Knowledge Graph

Figure 1, it can lead to a definition that states *"a knowledge graph acquires and integrates information into an ontology and applies a reasoner to derive new knowledge"* (Ehrlinger L, & Wöß W, 2016).

Therefore, it can be inferred that, while the knowledge graph is not entirely a new technology, however the requirement for a unified, non-conflicting definition still remains.

Applications of Knowledge Graphs

Knowledge Graphs are being utilized widely in a multitude of fields, ranging from research to industry. Apart from the applications of Knowledge Graphs discussed previously, this section brings to attention some more use cases. Several innovative ideas can be observed. One proposal showcased that Intelligent Digital Twins architecture was enhanced with the use of Knowledge Graphs (Sahlab N & Rychkova I & Dandash O & Lachenal C & Schleich B, 2021; Su C & Han Y & Tang X & Jiang Q & Wang T & He Q, 2024). Upon summarizing the use of KGs in drug discovery, one article presented a case study on COVID-19 research, which has leveraged knowledge graphs to identify potential drug candidates for repurposing (MacLean F & 2021). The industrial application of Knowledge Graphs stems from their potential to enhance intelligent systems. Applications include the use of KGs for fault diagnosis (Liu Y & Zhang H & Li X & Wang J, 2021), the food and science industry (Min W & Liu Z & Wang C & Yang Y & Zhao J, 2022), process knowledge applications (Lv Y & Wang Z & Li J & Chen X, 2024), as well as the medical field (Qu J, 2022). Few use cases of KGs in the medical sector explored applications regarding the construction of a Knowledge Graph for the Traditional Medicine system of China (Traditional Chinese Medicine - TCM) (Zhang Y & Hao Y, 2024), a kidney stone disease-specific knowledge graph construction (Man J & Shi Y & Hu Z & Yang R & Huang Z & Zhou Y, 2024), proposal of an electronic health record-oriented knowledge graph system for collaborative clinical decision support (Shang Y & Zhang W & Liu P & Li J & Chen X, 2024). KGs are also said to be highly useful in the cases of automatic question answering (Zhang F & Zhang Y & Xu T, 2020). One study showcased a framework for the construction of a forestry policy knowledge graph upon addressing the lack of research on policy knowledge graph construction methods (Sun J & Luo Z, 2024). KGs have also been employed for chatbots, which in turn emphasizes their highly significant role in developing AI assistants (Rajabi E & George A N & Kumar K, 2024). The result of the lack of guidance in agricultural remote sensing research, construction methods and applications of an agricultural remote sensing concept graph has also been discussed (Xu L & Zhang Y & Wang J & Li H, & Chen M, 2024). One study introduced an intelligent adaptive learning model, which is enabled by a knowledge graph and BiLSTM-CRF, showcasing its potential applications (Zhu Y, 2024).

Challenges of Knowledge Graphs

While KGs have become a popular tool to represent and organize complex information in various domains, their construction and management present several challenges that researchers are actively working to address.

Studies have highlighted the importance of correctness and completeness in dynamic knowledge graphs (DKGs) and discussed the impact of factors such as data sources, graph construction models, and evaluation methods on the accuracy of dependent applications (Farghaly M & Ali H & El-Meligy M & Youssef A, 2024). The challenges and progress have been discussed in graph domain adaptation (GDA), which combines graph representation learning and domain adaptation to facilitate transfer learning on graphs (Shi B & Zhang T & Li X & Wang Y, 2024). Studies also emphasize the challenges in representing heterogeneous graphs and the need for effective measurement, analysis, and mining techniques (Wang X & Li J & Zhao R & Zhou Q, 2024). Refinement methods include a proposal of data management techniques to scale up the creation of KGs (Iglesias E & Fernández J D & Priyatna F & Corcho O, 2024), introduction of new types of representational entities in knowledge graphs (Vogt L, 2024), accurate retrieval from textual KGs for answering complex real-world questions (Huang J & Zhao T & Liu Y & Wang X, 2024), and a novel graph neural network model (Ding W & Cherukumalli M & Chikoti S & Chaudhri V K, 2024) to name a few.

From the time Google released the google knowledge graph, research on KGs has flourished, achieving advancements in tasks like searching, entity resolution, and link prediction, with applications spanning ecosystems from companies like eBay and Amazon to scientific fields such as biology, geology (Kejriwal M, 2022). As research in this field progresses, novel methodologies and applications are emerging to enhance the capabilities of knowledge graphs (Ji S & Pan S & Cambria E & Marttinen P & Yu P S, 2021).

Natural Language Processing (NLP)

NLP, a branch of artificial intelligence (AI) and linguistics, is dedicated to empowering computers to process, interpret, generate, and engage with human language. NLP acts as a bridge between human language (aka, natural language) and machine understanding, thereby making it possible for computers to process textual and speech data and analyze large volumes of it. The core tasks of NLP include text understanding (breaking text into smaller units, identifying certain entities, and understanding grammatical roles and relationships), text analysis (sentiment analysis, identifying main topics of discussion, and classifying text into predefined labels), language generation (text summarization, language translation, question answering, and chatbots and conversational AI), and speech and language interaction (spoken language converted to text, and text converted to spoken language).

Significant advancements can be seen in the domain of NLP, with researchers exploring various techniques and applications within the field. Highly informative studies can be found that focus on showcasing the potential of randomized algorithms in NLP tasks (Ravichandran D & Pantel P & Hovy E, 2005), the importance of domain adaptation in NLP (Jiang J & Zhai C, 2007), models for understanding and processing metaphorical language in NLP (Shutova E, 2010), the importance of statistical analysis in NLP research (Berg-Kirkpatrick T & Burkett D & Klein D, 2012), the significance of interpretability in neural NLP systems (Li J & Chen X & Hovy E & Jurafsky D, 2015) and transferability of neural networks (Mou L & Meng Z & Yan R & Li G & Xu Y & Zhang L & Jin Z, 2016), and techniques for clinical text processing in diverse linguistic contexts (Névéol A & Dalianis H & Velupillai S & Savova G & Zweigenbaum P, 2018).

Applications of NLP

The application of NLP is widely explored in various fields, showcasing its versatility and effectiveness in different domains. The application of NLP in social media data analysis was showcased by introducing a technique for sentiment analysis in tweets (Mahzari M & Alshammari A S & Mehdizadeh A & Ghaffari M, 2024). The potential of NLP in synthesizing existing research and identifying future research directions has also been highlighted (Ye J & Wang Y & Liu X & Zhang H, 2024). The incorporation of NLP techniques in medical education program evaluation has been explored (Costa-Dookhan K A & Jones T & Patel M & Smith R, 2024). Additionally, research on NLP methods is applied in dermatology (Paganelli A & Rossi M & Bianchi L & Ferrara G, 2024) and psoriasis (Shapiro J & Klein T & Williams P & Chen H, 2024), thus showcasing the diverse applications of NLP in healthcare settings. One study emphasized the importance of advancing rheumatology with NLP techniques to improve the detection and diagnosis of disease, along with patient management, underscoring the need for targeted research to fully realize NLP's potential in clinical practice (Omar M & Ali N & Patterson C, 2024). Another study validated the use of an oncology NLP model in extracting clinical insights from non-small cell lung cancer data, demonstrating the reliability and generalizability of NLP models in supporting research studies and clinical trials (Kenney R C & Shah P & Gupta L & Brown D, 2024).

Challenges of NLP

Natural Language Processing has seen significant advancements in various domains, but it also faces several challenges that researchers and practitioners need to address. One of the challenges highlighted in the literature is the processing of highly unstructured clinical notes lacking proper grammar and punctuation in the healthcare domain, which makes data processing extremely difficult (Mekhtieva R L & Thompson B & Wang X & White S, 2024). Another challenge in NLP is the integration of generative AI tools into enterprise-critical software systems (Ghaisas S & Singhal A, 2024). Regarding text classification, particularly in multi-label text classification (MLTC), learning effective representations remains a significant challenge in NLP (Audibert A & Gauffre A & Amini M R, 2024). One study highlighted the ethical implications surrounding the use of generative AI tools in NLP in a comparative analysis of various AI tools, talking about their pros and cons and the regulatory challenges faced (Iorliam A & Ingio J A, 2024). Challenges regarding language are quite common regarding contextual words, synonyms, homonyms, sarcasm and irony, sentences with ambiguity, phrases that sound informal, expressions, idioms, and culture-specific lingo (Khurana D & Koli A & Khatter K & Singh S, 2023).

RELATED WORK ON KNOWLEDGE GRAPHS IN NLP

This section examines the Knowledge Graphs and its integration with natural language processing techniques. Knowledge Graphs have become an essential component in Natural Language Processing tasks, providing a structured representation of information and relationships within text data. They allow for the integration of structured knowledge with data-driven models, brought about by the blending of neural models and knowledge graphs for enhanced representations beyond individual approaches (Gomez-Perez J M & Denaux R & Garcia-Silva A, 2020).

Fundamental NLP tasks such as lexical analysis, syntactic analysis, and semantic understanding and advanced technologies like word embedding are covered before delving into the complete process of knowledge graph construction (Wang Z, 2024). NLP can help unlock the full potential of scientific knowledge graphs by connecting unstructured text with structured data, paving the way for more intelligent and automated research workflows (Quevedo X & Chicaiza J, 2023). A survey conducted on KGs in NLP noted that while studies spanning various domains explored up-and-coming topics such as knowledge graph embedding or augmented language model, a lack of secondary research and evaluations in practice still exists (Schneider P & Schopf T & Vladika J & Galkin M & Simperl E & Matthes F, 2022).

Graph Neural Networks (GNNs) used for knowledge graph rewiring and document classification in NLP, in order to extract complex patterns within text data and relationships present in those data entities have been explored (Romanova A, 2024). One study focused on predicting treatment relations between biomedical entities using semantic patterns over biomedical knowledge graphs to identify new possible treatment options for medical conditions (Bakal, G., & Kavuluru R., 2015).

By studying related works, it was observed that a number of novel methods, tools, systems, and frameworks have been developed in recent years that leverage the integration of knowledge graphs and NLP. A two-stage annotation methodology to build a knowledge graph of NLP contributions directly from scholarly articles, highlighting the limitations of structuring contributions compared to other STEM fields, was developed (D'souza J & Auer S, 2021). As a way of improving semantic analysis in biomedical NLP applications, a method is proposed that combines two models: one of knowledge graph-based language and another nearest-neighbor (Naseem U & Khan M A & Hussain M & Kim J, 2023). As autonomous NLP methods may lack accuracy in the creation of high-quality knowledge graphs, a study introduced a methodology named 'TinyGenius', which validates NLP-extracted knowledge statements using microtask crowdsourcing (Oelen A & Stocker M & Auer S, 2022). A group-specific approach to NLP for hate speech detection was utilized, leveraging a knowledge graph of antisemitic history and language to improve model performance (Halevy K, 2023). One study developed NLP-KG, a feature-rich system for exploratory searches of scientific literature in NLP fields, aiming to support researchers in navigating unfamiliar research areas (Schopf T & Matthes F, 2024). Another study introduced the ARCH system, which constructs a large-scale knowledge graph by analyzing and aggregating codified narrative health records, providing valuable clinical insights for research and clinical care (Gan Z & Wang T & Li J & Zhang H, 2023). Another such tool presented is the web-based CleanGraph, designed for human-in-the-loop refinement and completion of knowledge graphs (Bikaun T & Stewart M & Liu W, 2024). A framework named Graphusion has been presented for constructing knowledge graphs from free text, emphasizing the importance of KGs in artificial intelligence applications and downstream tasks like Question Answering systems (Yang R & Yang B & Ouyang S & She T & Feng A & Jiang Y & Li I, 2024). A Political Experts through Knowledge Graph Integration (PEG) framework to address challenges in aggregating and comprehending political information has been proposed, highlighting the importance of integrating local and global knowledge (Mou X & Li Z & Lyu H & Luo J & Wei Z, 2024). One study focused on Community Knowledge Graph Abstraction for Enhanced Link Prediction, specifically studying the PubMed Knowledge Graph (Zhao Y & Li M & Sun J & Xu B, 2024). An automatic knowledge graph constructed from literature and NLP-based reasoning network framework and ontology was created, emphasizing structured knowledge representation in NLP tasks (Chen H & Luo X, 2019). A knowledge extraction framework, a 2-phase framework, was introduced, decoupling knowledge extraction from English texts with competitive precision and recall rates (Corcoglioniti F & Rospocher M & Aprosio A P, 2016). As a way of integrating extra knowledge into word embedding models for biomedical NLP tasks, GCBOW and GSkip-gram models were proposed, which incorporate graphs into CBOW and Skip-gram models through graph regularization (Ling Y & Xu H & Liu Y & Zhao Y, 2017). SCICERO is a novel approach for automatically generating Scientific Knowledge Graphs in Computer Science by extracting entities and relationships from research papers using NLP techniques, deep learning models, and ontology-based validation (Dessí D & Osborne F & Atzori M & Motta E, 2022). A study presented a novel method that enhances NLP models by integrating knowledge graphs, using attention mechanisms and convolution-based encoding to improve performance, reduce reliance on labeled datasets, and significantly boost text classification and natural language inference accuracy (Annervaz K M & Chowdhury S B R & Dukkipati A, 2018). A pipeline was developed to transform scientific publications to a structured scientific knowledge graph, by integration of advanced NLP tools and machine learning techniques, merging their results to represent detailed knowledge in the Semantic Web domain, with evaluations demonstrating that tool integration improves the quality and performance of the resulting graphs (Dessì D & Osborne F & Reforgiato Recupero D & Motta E, 2021).

Studies have also identified the limiting factors of the current state of this domain and proposed novel solutions for the identified issues. The Geography-Graph Pre-trained model (G2PTL) addresses existing NLP model limitations in geographic knowledge to enhance geospatial tasks (Wu L & Li H & Zhang P & Zhao W, 2024). The GLAME model was proposed to address challenges in updating knowledge in LLMs, using knowledge graphs to enhance editing and improve post-edit generalization (Zhang M & Ye X & Liu Q & Ren P & Wu S & Chen Z, 2024). A Graph Recurrent Network (GRN) was introduced in order to effectively tackle graphical NLP challenges (Song L, 2019). One study highlighted the limitations in Sanskrit NLP for automated knowledge base construction, leading to manual annotation efforts for knowledge graph creation (Terdalkar H & Bhattacharya A & Dubey M & Singh B N, 2022).

A review of advances in named-entity extraction as a key task for transforming natural language texts into knowledge graphs highlights the need for disambiguation, linking, and joint learning approaches integrated with modern NER techniques while emphasizing challenges such as the lack of evaluation standards and the shift towards deep-learning-based, end-to-end systems that analyze mentions and entities in context (Al-Moslmi T & Ocaña M G & Opdahl A L & Veres C, 2020).

Overall, it can be said that the integration of knowledge graphs with NLP models has shown immense promise in enhancing semantic analysis, knowledge representation, and information retrieval in various research domains. However, more and more research is required to explore the full capabilities of knowledge graphs for NLP applications.

Future Directions

Future research in KGs for NLP is expected to focus on deeper integration of KGs with large language models (LLMs), which will enhance accuracy, reasoning and explainability. Advances in automated KG construction using self-supervised and deep learning techniques will improve scalability and reduce manual effort in entity and relation extraction. The emergence of multimodal knowledge graphs, incorporating text, images, and structured data, will enable richer context understanding in NLP applications. Additionally, KG-based reasoning will play a crucial role in improving interpretability and few-shot learning capabilities. Personalization and context-aware NLP systems will benefit from dynamically adaptive KGs, enhancing applications such as recommendation systems and conversational AI. Standardization efforts in KG evaluation metrics and dataset benchmarking will be necessary to ensure interoperability and reliability. Lastly, addressing ethical concerns, including bias, fairness, and data privacy, will be critical in developing responsible KG-driven NLP models for real-world deployment.

Acknowledgment

All the authors of this paper express their sincere thanks to their colleagues and paper reviewers for their timely and insightful comments, which have contributed to enhancing this comprehensive analysis.

Conflict of Interest

all the authors of this paper mention that they have no conflicts of interest concerning this work.

References

Al-Moslmi, T., Ocaña, M. G., Opdahl, A. L., & Veres, C. (2020). Named entity extraction for knowledge graphs: A literature overview. *IEEE Access*, *8*, 32862-32881..

Annervaz, K. M., Chowdhury, S. B. R., & Dukkipati, A. (2018). Learning beyond datasets: Knowledge graph augmented neural networks for natural language processing. *arXiv preprint arXiv:1802.05930*.

Audibert, A., Gauffre, A., & Amini, M. R. (2024, August). Exploring contrastive learning for long-tailed multi-label text classification. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 245-261). Cham: Springer Nature Switzerland.

Bakal, G., & Kavuluru, R. (2015, December). Predicting treatment relations with semantic patterns over biomedical knowledge graphs. In *International Conference on Mining Intelligence* *and Knowledge Exploration* (pp. 586-596). Cham: Springer International Publishing.

Berg-Kirkpatrick, T., Burkett, D., & Klein, D. (2012, July). An empirical investigation of statistical significance in NLP. In *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning* (pp. 995-1005).

Bikaun, T., Stewart, M., & Liu, W. (2024). CleanGraph: Human-in-the-loop Knowledge Graph Refinement and Completion. *arXiv preprint arXiv:2405.03932*.

Blumauer, A. (2014). From taxonomies over ontologies to knowledge graphs. *Semantic Web Blog*. https://blog.semanticweb.at/2014/07/15/from-taxonomies-over-ontologies-to-knowledge-graphs.

Chen, H., & Luo, X. (2019). An automatic literature knowledge graph and reasoning network modeling framework based on ontology and natural language processing. *Advanced Engineering Informatics*, *42*, 100959.

Corcoglioniti, F., Rospocher, M., & Aprosio, A. P. (2016, April). A 2-phase frame-based knowledge extraction framework. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing* (pp. 354-361).

Costa-Dookhan, K. A., Maslej, M. M., Donner, K., Islam, F., Sockalingam, S., & Thakur, A. (2024). Twelve tips for Natural Language Processing in medical education program evaluation. *Medical Teacher*, *46*(9), 1147-1151.

Dessí, D., Osborne, F., Recupero, D. R., Buscaldi, D., & Motta, E. (2022). SCICERO: A deep learning and NLP approach for generating scientific knowledge graphs in the computer science domain. *Knowledge-Based Systems*, *258*, 109945.

Dessì, D., Osborne, F., Recupero, D. R., Buscaldi, D., & Motta, E. (2021). Generating knowledge graphs by employing natural language processing and machine learning techniques within the scholarly domain. *Future Generation Computer Systems*, *116*, 253-264.

Ding, W., Cherukumalli, M., Chikoti, S., & Chaudhri, V. K. (2024, July). JPEC: A Novel Graph Neural Network for Competitor Retrieval in Financial Knowledge Graphs. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 2755-2759).

D'souza, J., & Auer, S. (2021). Sentence, phrase, and triple annotations to build a knowledge graph of natural language processing contributions-A trial dataset.

Ehrlinger, L., & Wöß, W. (2016). Towards a definition of knowledge graphs. *SEMANTiCS (Posters, Demos, SuCCESS)*, *48*(1-4), 2.

Färber, M., Bartscherer, F., Menne, C., & Rettinger, A. (2017). Linked data quality of dbpedia, freebase, opencyc, wikidata, and yago. *Semantic Web*, *9*(1), 77-129.

Farghaly, M., Mounir, M., Aref, M., & Moussa, S. M. (2024). Investigating the challenges and prospects of construction models for dynamic knowledge graphs. *IEEE Access*, *12*, 40973-40988.

Gan, Z., Zhou, D., Rush, E., Panickan, V. A., Ho, Y. L., Ostrouchovm, G., ... & Lu, J. (2025). Arch: Large-scale knowledge graph via aggregated narrative codified health records analysis. *Journal of Biomedical Informatics*, 104761.

Ghaisas, S., & Singhal, A. (2025). Dealing with Data for RE: Mitigating Challenges while using NLP and Generative AI. In *Handbook on Natural Language Processing for Requirements Engineering* (pp. 457-486). Cham: Springer Nature Switzerland.

Gomez-Perez, J. M., Denaux, R., & Garcia-Silva, A. (2020). *A Practical guide to hybrid natural language processing: Combining neural models and knowledge graphs for NLP* (Vol. 281). Springer.

Halevy, K. (2023). A group-specific approach to nlp for hate speech detection. *arXiv preprint arXiv:2304.11223*.

Hitzler, P. (2021). A review of the semantic web field. *Communications of the ACM, 64*(2), 76-83.

Hogan, A., Blomqvist, E., et.al (2021). Knowledge graphs. *ACM Computing Surveys, 54*(4), 1-37.

Huang, J., Li, M., Yao, Z., Yang, Z., Xiao, Y., Ouyang, F., ... & Yu, H. (2024). Ritek: A dataset for large language models complex reasoning over textual knowledge graphs. *arXiv preprint arXiv:2410.13987*.

Iglesias, E., Vidal, M. E., Collarana, D., & Chaves-Fraga, D. (2025). Empowering the SDM-RDFizer tool for scaling up to complex knowledge graph creation pipelines. *Semantic Web*, *16*(2), SW-243580.

Iorliam, A., & Ingio, J. A. (2024). A comparative analysis of generative artificial intelligence tools for natural language processing. *Journal of Computing Theories and Applications*, *1*(3), 311-325.

Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P. S. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(2), 494-514.

Jiang, J., & Zhai, C. (2007). Instance weighting for domain adaptation in NLP. ACL.

Kejriwal, M. (2022). Knowledge graphs: A practical review of the research landscape. Information, 13(4), 161.

Kenney, R. C., Chen, X., Shintani, K., Gagnon, C., Liu, J., DaCosta Byfield, S., ... & Currie, A. M. (2024). Validation of Non–Small Cell Lung Cancer Clinical Insights Using a Generalized Oncology Natural Language Processing Model. *JCO Clinical Cancer Informatics*, *8*, e2300099.

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. *Multimedia tools and applications*, *82*(3), 3713-3744.

Kroetsch, M., & Weikum, G. (2016). Journal of Web Semantics: Special Issue on relationships between entities. *Knowledge Graphs*.

Li, J., Chen, X., Hovy, E., & Jurafsky, D. (2015). Visualizing and understanding neural models in NLP. *arXiv preprint arXiv:1506.01066*.

Ling, Y., An, Y., Liu, M., Hasan, S. A., Fan, Y., & Hu, X. (2017, May). Integrating extra knowledge into word embedding models for biomedical NLP tasks. In *2017 International joint conference on neural networks (IJCNN)* (pp. 968-975). IEEE.

Liu, Y., Ma, Y., Mao, Z., Jiang, B., Liu, J., & Bian, J. (2021, October). TD-GAT: Graph neural network for fault diagnosis knowledge graph. In *2021 China Automation Congress (CAC)* (pp. 1734-1739). IEEE.

Lv, Y., Wang, P., Ji, G., & Wang, L. (2024, June). Process hyper-relation knowledge graph construction and application. In *Journal of Physics: Conference Series* (Vol. 2787, No. 1, p. 012063). IOP Publishing.

Man, J., Shi, Y., Hu, Z., Yang, R., Huang, Z., & Zhou, Y. (2024). KSDKG: construction and application of knowledge graph for kidney stone disease based on biomedical literature and public

databases. *Health information science and systems*, *12*(1), 54.
MacLean, F. (2021). Knowledge graphs and their applications in drug discovery. *Expert opinion on drug discovery*, *16*(9), 1057-1069.
Mahzari, M., Hashim, A. H. A., Saeed, K. M. O., & Mokhtar, M. M. O. (2024). Two-Person Intuitionistic Neutrosophic Soft Games with Harris Hawks Optimizer based Tweets Classification on NLP Applications. *Full Length Article*, *24*(1), 314-14.
Mekhtieva, R. L., Forbes, B., Alrajeh, D., Delaney, B., & Russo, A. (2024, January). RECAP-KG: Mining Knowledge Graphs from Raw Primary Care Physician Notes for Remote COVID-19 Assessment in Primary Care. In *AMIA Annual Symposium Proceedings* (Vol. 2023, p. 1145).
Min, W., Liu, C., Xu, L., & Jiang, S. (2022). Applications of knowledge graphs for food science and industry. *Patterns*, *3*(5).
Mou, L., Meng, Z., Yan, R., Li, G., Xu, Y., Zhang, L., & Jin, Z. (2016). How transferable are neural networks in NLP applications. *arXiv preprint arXiv:1603.06111*.
Mou, X., Li, Z., Lyu, H., Luo, J., & Wei, Z. (2024, May). Unifying local and global knowledge: Empowering large language models as political experts with knowledge graphs. In *Proceedings of the ACM Web Conference 2024* (pp. 2603-2614).
Naseem, U., Thapa, S., Zhang, Q., Hu, L., Masood, A., & Nasim, M. (2023, July). Reducing knowledge noise for improved semantic analysis in biomedical natural language processing applications. In *Proceedings of the 5th Clinical Natural Language Processing Workshop* (pp. 272-277).
Névéol, A., Dalianis, H., Velupillai, S., Savova, G., & Zweigenbaum, P. (2018). Clinical natural language processing in languages other than English: opportunities and challenges. *Journal of biomedical semantics*, *9*, 1-13.
Noy, N., Gao, Y., Jain, A., Narayanan, A., Patterson, A., & Taylor, J. (2019). Industry-scale knowledge graphs: Lessonsand challenges: Five diverse technology companies show how it's done. *Queue, 17*(2), 48-75.
Oelen, A., Stocker, M., & Auer, S. (2022). TinyGenius: Intertwining natural language processing with microtask crowdsourcing for scholarly knowledge graph creation. *Proceedings of the 22nd ACM/IEEE Joint Conference on Digital Libraries* (JCDL 2022).
Omar, M., Ali, N., & Patterson, C. (2024). Advancing rheumatology with natural language processing: Insights and prospects from a systematic review. *Rheumatology Advances in Practice*, 8(4), rkae120.
Paganelli, A., Rossi, M., Bianchi, L., & Ferrara, G. (2024). Natural language processing in dermatology: A systematic literature review and state of the art. *Journal of the European Academy of Dermatology and Venereology*.
Paulheim, H. (2016). Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web, 8*(3), 489-508.
Qu, J. (2022). A review on the application of knowledge graph technology in the medical field. Scientific Programming, 2022(1), 3212370.
Quevedo, X., & Chicaiza, J. (2023). Using NLP to enrich scientific knowledge graphs: A case study to find similar papers. *Proceedings of the 15th International Conference on Agents and Artificial Intelligence* (ICAART), 2.
Rajabi, E., George, A. N., & Kumar, K. (2024). The role of knowledge graphs in chatbots. *The Electronic Library*.
Ravichandran, D., Pantel, P., & Hovy, E. (2005). Randomized algorithms and NLP: Using locality sensitive hash functions for high-speed noun clustering. *In Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics* (ACL'05) (pp. 622-629).
Romanova, A. (2024). Enhancing NLP through GNN-driven knowledge graph rewiring and document classification. *2024 35th Conference of Open Innovations Association* (FRUCT). IEEE.
Sahlab, N., Rychkova, I., Dandash, O., Lachenal, C., & Schleich, B. (2021). Knowledge graphs as enhancers of intelligent digital twins. *2021 4th IEEE International Conference on Industrial Cyber-Physical Systems* (ICPS). IEEE.
Schneider, P., Schopf, T., Vladika, J., Galkin, M., Simperl, E., & Matthes, F. (2022). A decade of knowledge graphs in natural language processing: A survey. *arXiv preprint* arXiv:2210.00105.
Schopf, T., & Matthes, F. (2024). Nlp-kg: A system for exploratory search of scientific literature in natural language processing. *arXiv preprint arXiv*:2406.15294.
Shang, Y., Zhang, W., Liu, P., Li, J., & Chen, X. (2024). Electronic health record–oriented knowledge graph system for collaborative clinical decision support using multicenter fragmented medical data: Design and application study. *Journal of Medical Internet Research*, 26, e54263.
Shapiro, J., Klein, T., Williams, P., & Chen, H. (2024). Application of an NLP AI tool in psoriasis: A cross-sectional comparative study on identifying affected areas in patients' data. Clinics in Dermatology.
Shi, B., Zhang, T., Li, X., & Wang, Y. (2024). Graph domain adaptation: Challenges, progress, and prospects. *arXiv Preprint, arXiv*:2402.00904. https://arxiv.org/abs/2402.00904
Shutova, E. (2010). Models of metaphor in NLP. Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics.
Singhal, A. (2012). Introducing the knowledge graph: things, not strings. Officials google blog, 5(16), 3.
Song, L. (2019). Tackling graphical NLP problems with graph recurrent networks. University of Rochester
Su, C., Han, Y., Tang, X., Jiang, Q., Wang, T., & He, Q. (2024). Knowledge-based digital twin system: Using a knowledge-driven approach for manufacturing process modeling. *Computers in Industry*. https://doi.org/10.1016/j.compind.2024.104101
Sun, J., & Luo, Z. (2024). ForPKG-1.0: A framework for constructing forestry policy knowledge graph and application analysis. *arXiv Preprint*, arXiv:2411.11090. https://arxiv.org/abs/2411.11090
Terdalkar, H., Bhattacharya, A., Dubey, M., & Singh, B. N. (2022). Semantic Annotation and Querying Framework based on Semi-structured Ayurvedic Text. *arXiv preprint* arXiv:2202.00216.
Vogt, L. (2024). Semantic Units: Increasing Expressivity and Simplicity of Formal Representations of Data and Knowledge in Knowledge Graphs. *arXiv preprint* arXiv:2407.10720.
Wang, X., Li, J., Zhao, R., & Zhou, Q. (2024). AceMap: Knowledge discovery through academic graph. *arXiv Preprint*, arXiv:2403.02576. https://arxiv.org/abs/2403.02576
Wang, Z. (2024). Information extraction and knowledge map construction based on natural language processing. *Frontiers in Computing and Intelligent Systems*, 7(2), 47–49.
Wu, L., Li, H., Zhang, P., & Zhao, W. (2024). G2PTL: A Geography-Graph Pre-trained Model. *Proceedings of the 33rd ACM*

International Conference on Information and Knowledge Management (CIKM 2024).

Xu, L., Zhang, Y., Wang, J., Li, H., & Chen, M. (2024). Concept graph construction and applied research of agricultural remote sensing. *International Journal of Remote Sensing*, 45(13), 4428-4448.

Yang, R., Yang, B., Ouyang, S., She, T., Feng, A., Jiang, Y., ... & Li, I. (2024). Graphusion: Leveraging large language models for scientific knowledge graph fusion and construction in nlp education. *arXiv preprint arXiv*:2407.10794.

Ye, J., Wang, Y., Liu, X., & Zhang, H. (2024). Development and application of natural language processing on unstructured data in hypertension: A scoping review. medRxiv Preprint.

Zhang, F., Zhang, Y., & Xu, T. (2020). A review of knowledge graph technology in the field of automatic question answering. *2020 International Signal Processing, Communications and Engineering Management Conference* (ISPCEM). IEEE.

Zhang, M., Ye, X., Liu, Q., Ren, P., Wu, S., & Chen, Z. (2024). Knowledge graph enhanced large language model editing. *arXiv preprint* arXiv:2402.13593.

Zhang, Y., & Hao, Y. (2024). Traditional Chinese medicine knowledge graph construction based on large language models. *Electronics*, 13(7), 1395.

Zhao, Y., Li, M., Sun, J., & Xu, B. (2024). Community knowledge graph abstraction for enhanced link prediction: A study on PubMed knowledge graph. *Journal of Biomedical Informatics*, 158, 104725.

Zhu, Y. (2024). A knowledge graph and BiLSTM-CRF-enabled intelligent adaptive learning model and its potential application. *Alexandria Engineering Journal*, 91, 305-320.