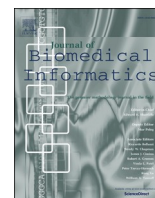




Contents lists available at ScienceDirect

# Journal of Biomedical Informatics

journal homepage: [www.elsevier.com/locate/yjbin](http://www.elsevier.com/locate/yjbin)

## Editorial

### Semantics-enabled biomedical literature analytics



#### 1. Introduction

Due to the large size and exponential growth of the number of scientific articles published in the biomedical domain, obtaining the most relevant articles to a topic of interest, identifying reliable and scientifically sound studies, extracting salient biomedical information from texts, and synthesizing knowledge from different studies present significant challenges. While traditional term-based information analysis and machine learning techniques can be employed for literature search, information extraction, and knowledge integration, such approaches lack an effective mechanism for analyzing scientific articles based on their semantics, which is often ambiguous and contextual. In the biomedical domain, knowledge representation and semantics-enabled techniques have shown the potential for systematic retrieval, curation, organization, and interpretation of content in ways that relates well to human understanding [1–5].

Here, we curate a special topic issue around the theme of *Semantics-enabled Biomedical Literature Analytics*. This Special Issue aims to highlight the development of novel informatics approaches for retrieval, indexing, and analysis of biomedical literature, focusing on semantics-based techniques. We place specific attention on methods that allow for the construction, analysis, and integration of biomedical knowledge bases (KBs), with the ultimate objective of employing such KBs for improving search performance over biomedical literature, enhancing biomedical information extraction, and literature-based knowledge discovery.

Biomedical text analysis, information retrieval (IR), and information extraction (IE), as well as biomedical KBs, ontologies, and vocabularies have been the topic of a number of special issues in recent years. A special issue in this journal [6] focused on novel approaches to collect, integrate, and analyze COVID-19 relevant data, which included scientific publications. A *Journal of the American Medical Informatics Association (JAMIA)* special issue [7] focused on the Unified Medical Language System (UMLS) [8], as an important source of structured biomedical semantics. The issue published by *BMC Medical Informatics and Decision Making* [9] is dedicated to biomedical terminologies and ontologies in general and their quality assurance. A special issue in *ACM Transactions on Computing for Healthcare* [10] covered techniques for natural language processing (NLP) of biomedical text. While the current Special Issue is synergistic with these collections, it has a distinct focus: it is dedicated to covering emerging approaches and strategies for using biomedical knowledge representation, ontologies, and other semantic resources in biomedical literature search and analytics.

We solicited submissions describing novel methodological results on the topics of biomedical informatics, knowledge representation/

ontologies, IE, NLP, artificial intelligence/machine learning, data mining, and other related areas. We accepted 15 papers to be included in this Special Issue.

#### 2. Research themes

We organize the 15 papers included as part of this Special Issue under the four broad themes of (i) Literature-based Discovery, (ii) Automated KB Construction, (iii) Knowledge-augmented Biomedical NLP, and (iv) Literature Search and IR. The accepted papers, organized by their primary theme, are listed in Table 1. We briefly describe each theme and representative papers below.

##### 2.1. Theme I: Literature-based discovery

Papers on this theme incorporate structured data from KBs and knowledge graphs to facilitate literature-based discovery. Cuffy and McInnes [11] explore ways of meaningfully connecting non-interacting works through known biomedical concepts, while Launer-Wachs et al. [12] use rapidly constructed structured knowledge resources from literature to support hypothesis generation. Pu et al. [13] introduce a framework for literature-based discovery for Alzheimer's Disease.

Cuffy and McInnes [11] introduce a deep learning neural network-based approach for literature-based discovery. The objective is to facilitate information sharing among non-interacting literature while extracting potentially meaningful information. More specifically, given two sets of biomedical concepts, the introduced neural model finds a set of concepts through which the former two concepts can be found related to each other.

Launer-Wachs et al. [12] describe a system for the ad-hoc construction of knowledge resources in the biomedical domain and present three case studies in which the system was used to rapidly construct KBs with varying degrees of verification. Their approach relies on an extractive search framework. The resulting KBs are proposed as a basis to support hypothesis verification for literature-based discovery. The authors suggest that the tool can be effectively used by biomedical researchers without expertise in bio-curation or NLP, and that the tool can support the generation of novel, plausible hypotheses, particularly using Swanson's ABC model [26]. Though focused on hypothesis generation, this paper uses rapid automated construction of KBs to facilitate this end goal.

Pu et al. [13] propose a framework for literature-based discovery based on graph embeddings and link prediction in the Alzheimer's Disease (AD) domain. They constructed a corpus from the AD publications published between 1977 and 2021 (titles and abstracts) and

<https://doi.org/10.1016/j.jbi.2024.104588>

Received 8 January 2024; Accepted 10 January 2024

Available online 19 January 2024

1532-0464/© 2024 Elsevier Inc. All rights reserved.

**Table 1**

Papers published as part of the Special Issue organized by theme.

<b>Theme I: Literature-based Discovery</b>	
Cuffy and McInnes [11]	Exploring a deep learning neural architecture for closed Literature-based discovery
Launer-Wachs et al. [12]	From centralized to ad-hoc knowledge base construction for hypotheses generation
Pu et al. [13]	Graph embedding-based link prediction for literature-based discovery in Alzheimer's Disease
<b>Theme II: Automated KB Construction</b>	
Boguslav et al. [14]	Creating an ignorance-base: Exploring known unknowns in the scientific literature
Dang et al. [15]	GENA: A knowledge graph for nutrition and mental health
Pérez-Pérez et al. [16]	A novel gluten knowledge base of potential biomedical and health-related interactions extracted from the literature: Using machine learning and graph analysis methodologies to reconstruct the bibliome
Sosa et al. [17]	Associating biological context with protein-protein interactions through text mining at PubMed scale
Taneja et al. [18]	Developing a Knowledge Graph for Pharmacokinetic Natural Product-Drug Interactions
<b>Theme III: Knowledge-augmented Biomedical NLP</b>	
Arabzadeh and Bagheri [19]	A self-supervised language model selection strategy for biomedical question answering
Badenes-Olmedo and Corcho [20]	Lessons learned to enable question answering on knowledge graphs extracted from scientific publications: A case study on the coronavirus literature
Lai et al. [21]	KEBLM: Knowledge-Enhanced Biomedical Language Models
Tsujimura et al. [22]	Large-scale neural biomedical entity linking with layer overwriting
<b>Theme IV: Literature Search and IR</b>	
Khader and Ensan [23]	Learning to rank query expansion terms for COVID-19 scholarly search
Lokker et al. [24]	Deep learning to refine the identification of high-quality clinical research articles from the biomedical literature: Performance evaluation
Nentidis et al. [25]	Large-scale investigation of weakly-supervised deep learning for the fine-grained semantic indexing of biomedical literature

annotated them with AD-specific information using the Neuropsychological Integrative Ontology [27] and PubTator Central [28]. After the construction of the knowledge graph, time slicing was used, 20 sets of training and testing sets were generated, and the performance of different models was evaluated. Their evaluation showed that Structural Deep Network Embedding (SDNE) model [29] consistently outperformed other methods and that the performance of the models improved significantly when the link prediction evaluation considers more distant future, reflecting the time required for knowledge accumulation.

## 2.2. Theme II: Automated KB construction

Rather than incorporating structured knowledge into language modeling, the papers on this theme focus on constructing KBs by extracting entities and relations from the biomedical literature. These KBs represent information in different domains such as the effect of gluten on health [16], nutrition and mental illness [15], and interactions between natural products and drugs [18]. Boguslav et al. [14] uniquely extract an *ignorance-base*, which focuses on identifying claims for which evidence is lacking in the literature. Lastly, while Sosa et al. [17] do not construct a new KB, they augment an existing protein-protein interaction database with additional information mined from the literature.

Boguslav et al. [14] extend prior work and report on the development of a type of knowledge base referred to as *ignorance-base*, which focuses not only on representing knowledge from the literature but also

providing additional annotations for claims that are ignorance statements, for which significant science is lacking. They develop an ignorance-base in the prenatal nutrition field to help graduate students search for thesis topics related to vitamin D's role in prenatal nutrition.

Dang et al. [15] introduce GENA, a knowledge graph that encodes relationships between food, biochemicals, and mental illnesses. The knowledge graph (entities and their relationships) is extracted from abstracts of PubMed articles. This paper uses a hybrid named entity recognition model to extract entities and then applies a syntax-based relation extraction model to detect relations between entities.

Pérez-Pérez et al. [16] describe the construction, application, and visualization of a KB for gluten health interactions developed from gluten-related literature using named entity recognition, relation extraction, and document classification methods followed by graph analysis. The authors also present case studies of applying the KB to analyze gluten-related health issues and the evidence on nutritional facts in the literature and suggest that the integration with social media can enable detection of gluten-related misinformation.

Sosa et al. [17] present a knowledge graph-context association task and methods for associating textual context with protein-protein interactions (PPIs) in KBs. The task is formulated as a classification task using syntactic, semantic, and discourse features. The approach is evaluated on highly precise silver standard corpora based on cell type or tissue contexts, which yields good performance, particularly for cell types. As a use case, they further incorporate cell type context into a protein-protein network for dengue fever.

Taneja et al. [18] propose NP-KG, a knowledge graph-based framework to investigate natural product-drug interactions (NPDIs). They construct a large-scale, heterogeneous knowledge graph from biomedical ontologies, linked data, and semantic relationships extracted from full-text biomedical publications using SemRep [30] and INDRA-REACH [31]. They evaluate NP-KG by investigating pharmacokinetic NPDIs involving two natural products, green tea and kratom, through KG path searches and *meta-path* discovery. They identify congruent, contradictory, and both congruent and contradictory information, when compared to manually curated ground truth knowledge.

## 2.3. Theme III: Knowledge-augmented biomedical NLP

Papers on this theme use information from a knowledge base to augment pre-trained language models (PLMs) [21] or improve the performance on specific NLP tasks such as question answering [19,20] or entity linking [22].

Arabzadeh and Bagheri [19] explore whether general-domain PLMs can be adapted for biomedical domain-specific tasks without fine-tuning. They introduce the idea of training a classifier that selects a general-domain language model from among a pool of models that is most likely to answer a domain-specific question correctly. Experiments are performed on the BioASQ dataset [32], and show that the proposed selection strategy can improve the performance of general-purpose language models, making them competitive with domain-specific models such as PubMedBERT [33].

Badenes-Olmedo and Corcho [20] first construct a knowledge graph of evidence on drugs, diseases, genes, and proteins by extracting entities and relationships from COVID-19, a dataset of COVID-19 scientific articles [34]. They then introduce a question-answering tool that uses the developed knowledge graph in addition to general-purpose knowledge graphs such as Wikidata and DBpedia to answer questions from the SimpleQuestions dataset [35].

Lai et al. [21] present a new technique for incorporating structured background knowledge into PLMs focusing on scientific literature. The technique uses adapter modules, that can extend PLMs with different types of domain knowledge. The authors demonstrate performance improvements on top of SciBERT [36] and BioBERT [37] models on biomedical entity linking, natural language inference, and question answering tasks.

Tsujimura et al. [22] describe a neural system for biomedical entity linking. The authors introduce three methodological enhancements compared to prior work: training data augmentation, layer overwriting, and a cosine similarity-based loss function. They demonstrate improvements on the n2c2 2019 Track 3 dataset [38], on which they achieve state-of-the-art results, and also provide experimental results on the MedMentions [39] and BioCreative VII Track 2 datasets [40], which shows significant improvements.

#### 2.4. Theme IV: Literature search and IR

This final theme incorporates structured semantic information to improve biomedical search and IR. The various works categorized under this theme aim to improve query expansion for COVID-19 search [23], enhance literature screening for clinical evidence broadly [24], and enable semantic indexing at a finer granularity than MeSH [25].

Khader and Ensan [23] present CQED, a Contextual Query Expansion framework based on domain knowledge to retrieve relevant COVID-19 articles. The framework first generates a set of query expansion terms based on domain knowledge and learns a learning-to-rank model that re-ranks the list of generated query expansion terms. The top-ranked terms are used to augment the original query when searching in PubMed. The authors conduct experiments on the TREC-COVID dataset [41], which includes queries on COVID-19 topics, demonstrating strong performance gains.

Lokker et al. [24] present a deep learning-based approach to literature screening for identifying articles with high-quality, clinically relevant evidence. They fine-tune several BERT variants on four annotated datasets and also ensemble the best-performing models on four datasets. A single model based on BioBERT [37] yields the best performance, improving specificity and maintaining high sensitivity in a prospective evaluation. The authors conclude that the model can increase the efficiency of literature surveillance and allow for faster dissemination of appraised research.

Nentidis et al. [25] propose a method for semantic indexing of biomedical literature at a finer granularity than that performed using MeSH (Medical Subject Headings) descriptors. Given the absence of annotated datasets, they utilize a trick in the MeSH evolution to automatically create datasets where such fine-grained labels can be identified (RetroBM). They combine weak supervision, dictionary-based heuristics, and deep learning models to index documents with fine-grained labels without ground-truth data (DBM). Their experimental results suggest that deep learning-based weakly-supervised fine-grained semantic indexing is promising for fine-grained access to the biomedical literature.

### 3. Summary

The collection of articles that appear in this Special Issue confirm several current trends. First, the use of PLMs, particularly domain-specific BERT variants, remains strong across various biomedical NLP tasks [14,17,19–25]. To some extent, these models implicitly capture semantic knowledge through large-scale pre-training on domain-specific corpora and often yield robust results. At the same time, their shortcomings in capturing nuanced domain knowledge through language modeling is acknowledged, motivating work that aims to incorporate semantic knowledge into PLMs explicitly to improve model performance on NLP tasks, using mechanisms such as adapters [21]. Knowledge fusion approaches show promise, given that biomedical domain is particularly rich in such structured knowledge.

Second, we observe that for practical applications of biomedical literature analysis, established tools and resources that provide biomedical semantic knowledge continue to be commonly used [13,16,18,25]. In addition, many systems use biomedical ontologies, and KBs, such as UMLS [8], MeSH, ChEBI [42]. In many cases, researchers incorporate multiple tools or ontologies into a single system.

**Table 2**  
PLMs, biomedical KBs and tools used by the studies in this issue.

Reference	PLMs	KBs	Tools
Cuffy and McInnes [11]	–	MeSH, CHEBI [42], NCBI Taxonomy [43], NCBI Gene [44]	PubTator [28]
Launer-Wachs et al. [12]	–	Any biomedical ontology	scispaCy [45], SPIKE [46]
Pu et al. [13]	–	Neuropsychological Integrative Ontology [27]	PubTator [28], ConceptMapper [47]
Boguslav et al. [14]	BERT [48], BioBERT [37]	CHEBI [42], NCBI Taxonomy [43], Cell Ontology [49], Gene Ontology [50], Protein Ontology [51], Sequence Ontology [52], Uberanatomy ontology [53]	PheKnowLator [54], DAVID [55]
Dang et al. [15]	–	CHEBI [42], Protein Ontology [51], DOID [56], Symptom Ontology [56], FMA [57], APADISORDERS [58], ASDITTO [59], FoodON [60]	scispaCy <sub>BC5CDR</sub> [45], PASMED [61], neo4j
Pérez-Pérez et al. [16]	–	MeSH, CHEBI [42], Symptom Ontology [56], FMA [57], FoodON [60], UniProt [62], Disease Ontology [63], KEGG [64], PharmGKB [65], DrugBank [66], NCI Thesaurus [67]	TMChem [68], LINNAEUS [69], DNORM [70], ABNER [71], OSCAR4 [72], jQuery, AmCharts, Angular.js, Cytoscape
Sosa et al. [17]	BioBERT [37], PubMedBERT [33]	MeSH	PubTator [28], GNormPlus [73], spaCy
Taneja et al. [18]	–	CHEBI [42], Cell Ontology [49], Gene Ontology [50], Protein Ontology [51], Sequence Ontology [52], Uberanatomy Ontology [53], Mondo Disease Ontology [74], Human Phenotype Ontology [75], Pathway Ontology [76], Cell Line Ontology [77]	SemRep [30], INDRAREACH [31], PheKnowLator [50]
Arabzadeh and Bagheri [19]	BERT [48], DistilBERT [78], RoBERTa [79], DistilRoBERTa, BioClinicalBERT [80], BioBERT [37], PubMedBERT [33]	–	–
Badenes-Olmedo and Corcho [20]	BioBERT [37]	MeSH, CHEBI [42], Gene Ontology [50], Protein Ontology [51], Human Disease Ontology [81], ICD-10 [82], PubChem [83], Anatomical Therapeutical Chemical Classification [84], OGC [85]	BioNER + BioNEN [86]
Lai et al. [21]	BioBERT [37], SciBERT. [87]	UMLS [8], PubChem [83], MSI [88]	–

(continued on next page)

Table 2 (continued)

Reference	PLMs	KBs	Tools
Tsujimura et al. [22]	BERT [48], SciBERT [87]	UMLS [8], MeSH	
Khader and Ensan [23]	BERT [48], UmlsBERT [89]	UMLS [8]	spaCy, scispaCy [45]
Lokker et al. [24]	BERT [48], BioBERT [37], BlueBERT [90], PubMedBERT [33]	–	–
Nentidis et al. [25]	PubMedBERT [33]	MeSH	MetaMap [91]

The tools and ontologies facilitate rapid application development and may mitigate the need for annotation; thus, we expect that they will continue to serve the research community in a wide range of analytics tasks, despite some shortcomings like accuracy and coverage, and the often non-trivial task of integrating these resources. Table 2 shows a complete overview of the use of PLMs, biomedical KB, ontologies, and tools across all accepted papers.

The submission deadline of the Special Issue more or less coincided with the release of ChatGPT by OpenAI and the unprecedented popularity of generative large language models (LLMs) trained using autoregressive language modeling [92–94]. While we received a couple of submissions that used such models, these were ultimately not accepted for this Special Issue. However, we note the increasing interest in using generative LLMs for in-context learning for many biomedical literature mining tasks [95,96]. Thus far, there is conflicting evidence on the effectiveness of such models on biomedical literature mining and IE tasks; some studies report that these models underperform fine-tuning [96,97], while others report competitive results [98]. At the same time, combining techniques such as fine-tuning open-source LLMs (e.g., Flan-T5 [99]) and chain-of-thought prompting [100] yields state-of-the-art performance on some datasets [101]. We anticipate that research on using generative LLMs for biomedical literature analysis will continue to make strides in the foreseeable future, including for the areas covered in this Special Issue. In fact, recently published work has demonstrated the promise of generative LLMs for literature-based discovery and KB construction (albeit in other scientific domains), while also pointing out their limited effectiveness compared to fine-tuned models for these tasks [102,103]. A natural direction could be to extend these methods to biomedical literature analysis tasks and to enhance their effectiveness by leveraging methods such as retrieval-augmented generation (RAG) [104,105].

Given their current tendency to hallucinate [106], it remains to be seen whether/when LLMs can fully substitute more robust and explainable methods in the near future. In this regard, we anticipate that combining LLMs with structured semantic knowledge (i.e., neuro-symbolic AI [107]) can be synergistic; hallucinations can be reduced or prevented by grounding LLM output in explicit knowledge, while LLMs can be used to address the problems of coverage and completeness that often diminish the utility of curated semantic resources [108].

#### CRedit authorship contribution statement

**Halil Kilicoglu:** Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing. **Faezeh Ensan:** Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing. **Bridget McInnes:** Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing. **Lucy Lu Wang:** Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

HK was partially supported by the National Institutes of Health grants R01LM014079 and R01LM014292. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funder had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

#### References

- [1] G. Tsatsaronis, G. Balikas, P. Malakasiotis, I. Partalas, M. Zschunke, M.R. Alvers, D. Weissenborn, A. Krithara, S. Petridis, D. Polychronopoulos, et al., An overview of the BIOASQ large-scale biomedical semantic indexing and question answering competition, *BMC Bioinf.* 16 (2015) 1–28.
- [2] H. Kilicoglu, D. Shin, M. Fiszman, G. Roseblat, T.C. Rindfleisch, SemMedDB: a PubMed-scale repository of biomedical semantic predications, *Bioinformatics* 28 (2012) 3158–3160.
- [3] D.N. Nicholson, C.S. Greene, Constructing knowledge graphs and their biomedical applications, *Computat. Struct. Biotechnol. J.* 18 (2020) 1414–1428.
- [4] R. Zhang, D. Hristovski, D. Schutte, A. Kastrin, M. Fiszman, H. Kilicoglu, Drug repurposing for COVID-19 via knowledge graph completion, *J. Biomed. Inform.* 115 (2021) 103696.
- [5] E. Wood, A.K. Glen, L.G. Kvarfordt, F. Womack, L. Acevedo, T.S. Yoon, C. Ma, V. Flores, M. Sinha, Y. Chodpathumwan, et al., RTX-KG2: a system for building a semantically standardized knowledge graph for translational biomedicine, *BMC Bioinf.* 23 (2022) 400.
- [6] H. Xu, D.L. Buckeridge, F. Wang, P. Tarczy-Hornoch, Novel informatics approaches to COVID-19 research: from methods to applications, *J. Biomed. Inform.* 129 (2022) 104028.
- [7] B.L. Humphreys, G. Del Fiol, H. Xu, The UMLS knowledge sources at 30: indispensable to current research and applications in biomedical informatics, *J. Am. Med. Inform. Assoc.* 27 (2020) 1499–1501.
- [8] O. Bodenreider, The Unified Medical Language System (UMLS): integrating biomedical terminology, *Nucleic Acids Res.* 32 (2004) D267–D270.
- [9] A. Agrawal, L. Cui, Quality assurance and enrichment of biological and biomedical ontologies and terminologies, *BMC Med. Inf. Decis. Making* 20 (2020) 1–4.
- [10] M.V. Devarakonda, E.M. Voorhees, Introduction to the Special Issue on Computational Methods for Biomedical NLP, *ACM Trans. Comput. Healthcare (HEALTH)* 3 (2022). URL: <https://doi.org/10.1145/3492302>. doi: 10.1145/3492302.
- [11] C. Cuffy, B.T. McInnes, Exploring a deep learning neural architecture for closed literature-based discovery, URL: *J. Biomed. Inform.* 143 (2023) 104362 <https://www.sciencedirect.com/science/article/pii/S1532046423000837> <https://doi.org/10.1016/j.jbi.2023.104362>.
- [12] S. Launer-Wachs, H. Taub-Tabib, J.T. Madem, O. Bar-Natan, Y. Goldberg, Y. Shamay, From centralized to ad-hoc knowledge base construction for hypotheses generation, URL: *J. Biomed. Inform.* 142 (2023) 104383 <https://www.sciencedirect.com/science/article/pii/S1532046423001041> <https://doi.org/10.1016/j.jbi.2023.104383>.
- [13] Y. Pu, D. Beck, K. Verspoor, Graph embedding-based link prediction for literature-based discovery in Alzheimer's Disease, URL: *J. Biomed. Inform.* 145 (2023) 104464 <https://www.sciencedirect.com/science/article/pii/S1532046423001855> <https://doi.org/10.1016/j.jbi.2023.104464>.
- [14] M.R. Boguslav, N.M. Salem, E.K. White, K.J. Sullivan, M. Bada, T.L. Hernandez, S. M. Leach, L.E. Hunter, Creating an ignorance-base: Exploring known unknowns in the scientific literature, *J. Biomed. Inform.* 143 (2023) 104405. URL: <https://www.sciencedirect.com/science/article/pii/S1532046423001260>, <https://doi.org/10.1016/j.jbi.2023.104405>.
- [15] L.D. Dang, U.T. Phan, N.T. Nguyen, GENA: A knowledge graph for nutrition and mental health, *J. Biomed. Inform.* 145 (2023) 104460. URL: <https://www.sciencedirect.com/science/article/pii/S1532046423001818>, <https://doi.org/10.1016/j.jbi.2023.104460>.
- [16] M. Pérez-Pérez, T. Ferreira, G. Igrejas, F. Fdez-Riverola, A novel gluten knowledge base of potential biomedical and health-related interactions extracted from the literature: Using machine learning and graph analysis methodologies to reconstruct the bibliome, *Journal of Biomedical Informatics* 143 (2023) 104398. URL: <https://www.sciencedirect.com/science/article/pii/S1532046423001193>. doi:<https://doi.org/10.1016/j.jbi.2023.104398>.
- [17] D.N. Sosa, R. Hintzen, B. Xiong, A. de Gioglio, J. Fauqueur, M. Davies, J. Lever, R. B. Altman, Associating biological context with protein-protein interactions through text mining at PubMed scale, URL: *J. Biomed. Inform.* 145 (2023) 104474 <https://www.sciencedirect.com/science/article/pii/S1532046423001958> <https://doi.org/10.1016/j.jbi.2023.104474>.

- [18] S.B. Taneja, T.J. Callahan, M.F. Paine, S.L. Kane-Gill, H. Kilicoglu, M. P. Joachimiak, R.D. Boyce, Developing a knowledge graph for pharmacokinetic natural product-drug interactions, *J. Biomed. Inform.* 140 (2023) 104341. URL: <https://www.sciencedirect.com/science/article/pii/S153204642300062X>, <https://doi.org/10.1016/j.jbi.2023.104341>.
- [19] N. Arabzadeh, E. Bagheri, A self-supervised language model selection strategy for biomedical question answering, *J. Biomed. Inform.* 146 (2023) 104486. URL: <https://www.sciencedirect.com/science/article/pii/S1532046423002071>, <https://doi.org/10.1016/j.jbi.2023.104486>.
- [20] C. Badenes-Olmedo, O. Corcho, Lessons learned to enable question answering on knowledge graphs extracted from scientific publications: A case study on the coronavirus literature, *J. Biomed. Inform.* 142 (2023) 104382. URL: <https://www.sciencedirect.com/science/article/pii/S153204642300103X>, <https://doi.org/10.1016/j.jbi.2023.104382>.
- [21] T.M. Lai, C. Zhai, H. Ji Keblm, Knowledge-Enhanced biomedical language models, URL: *J. Biomed. Inform.* 143 (2023) 104392 <https://www.sciencedirect.com/science/article/pii/S1532046423001132> <https://doi.org/10.1016/j.jbi.2023.104392>.
- [22] T. Tsujimura, M. Miwa, Y. Sasaki, Large-scale neural biomedical entity linking with layer overwriting, *J. Biomed. Inform.* 143 (2023) 104433. URL: <https://www.sciencedirect.com/science/article/pii/S1532046423001545>, <https://doi.org/10.1016/j.jbi.2023.104433>.
- [23] A. Khader, F. Ensan, Learning to rank query expansion terms for COVID-19 scholarly search, *J. Biomed. Inform.* 142 (2023) 104386. URL: <https://www.sciencedirect.com/science/article/pii/S1532046423001077>, <https://doi.org/10.1016/j.jbi.2023.104386>.
- [24] C. Lokker, E. Bagheri, W. Abdelkader, R. Parrish, M. Afzal, T. Navarro, C. Cotoi, F. Germini, L. Linkins, R.B. Haynes, L. Chu, A. Iorio, Deep learning to refine the identification of high-quality clinical research articles from the biomedical literature: performance evaluation, URL: *J. Biomed. Inform.* 142 (2023) 104384 <https://www.sciencedirect.com/science/article/pii/S1532046423001053> <https://doi.org/10.1016/j.jbi.2023.104384>.
- [25] A. Nentidis, T. Chatzopoulos, A. Krithara, G. Tsoumakas, G. Paliouras, Large-scale investigation of weakly-supervised deep learning for the fine-grained semantic indexing of biomedical literature, *J. Biomed. Inform.* 146 (2023) 104499. URL: <https://www.sciencedirect.com/science/article/pii/S1532046423002204>. doi: <https://doi.org/10.1016/j.jbi.2023.104499>.
- [26] D.R. Swanson, Fish oil, Reynaud's syndrome, and undiscovered public knowledge, *Perspect. Biol. Med.* 30 (1986) 7–18.
- [27] A. Gomez-Valades, R. Martinez-Tomas, M. Rincon, Integrative base ontology for the research analysis of Alzheimer's disease-related mild cognitive impairment, *Front. Neuroinf.* 15 (2021) 561691.
- [28] C.-H. Wei, A. Allot, R. Leaman, Z. Lu, PubTator central: automated concept annotation for biomedical full text articles, *Nucleic Acids Res.* 47 (2019) W587–W593.
- [29] D. Wang, P. Cui, W. Zhu, Structural deep network embedding, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016*, pp. 1225–1234.
- [30] H. Kilicoglu, G. Roseblat, M. Fiszman, D. Shin, Broad-coverage biomedical relation extraction with SemRep, *BMC Bioinf.* 21 (2020) 1–28.
- [31] M.A. Valenzuela-Escárcega, Ö. Babur, G. Hahn-Powell, D. Bell, T. Hicks, E. Noriega-Atala, X. Wang, M. Suredanu, E. Demir, C.T. Morrison, Large-scale automated machine reading discovers new cancer-driving mechanisms, *Database* (2018 (2018)) bay098.
- [32] A. Nentidis, A. Krithara, K. Bougiatiotis, M. Krallinger, C. Rodriguez-Penagos, M. Villegas, G. Paliouras, Overview of BioASQ 2020: The eighth BioASQ challenge on large-scale biomedical semantic indexing and question answering, in: *Experimental IR Meets Multilinguality, Multimodality, and Interaction: 11th International Conference of the CLEF Association, CLEF 2020, Thessaloniki, Greece, September 22–25, 2020, Proceedings 11*, Springer, 2020, pp. 194–214.
- [33] Y. Gu, R. Tinn, H. Cheng, M. Lucas, N. Usuyama, X. Liu, T. Naumann, J. Gao, H. Poon, Domain-specific language model pretraining for biomedical natural language processing, *ACM Trans. Comput. Healthcare (HEALTH)* 3 (2021) 1–23.
- [34] L. L. Wang, K. Lo, Y. Chandrasekhar, R. Reas, J. Yang, D. Burdick, D. Eide, K. Funk, Y. Katsis, R. M. Kinney, Y. Li, Z. Liu, W. Merrill, P. Mooney, D. A. Murdick, D. Rishi, J. Sheehan, Z. Shen, B. Stilson, A. Wade, K. Wang, N.X.R. Wang, C. Wilhelm, B. Xie, D.M. Raymond, D.S. Weld, O. Etzioni, S. Kohlmeier, *CORD-19: The COVID-19 open research dataset*, in: K. Verspoor, K.B. Cohen, M. Dredze, E. Ferrara, J. May, R. Munro, C. Paris, B. Wallace (Eds.), *Proceedings of the 1<sup>st</sup> Workshop on NLP for COVID-19 at ACL 2020, Association for Computational Linguistics, Online*, 2020. URL: <https://aclanthology.org/2020.nlpcovid19-acl.1>.
- [35] A. Bordès, N. Usunier, S. Chopra, J. Weston, Large-scale simple question answering with memory networks, *arXiv preprint arXiv:1506.02075* (2015).
- [36] I. Beltagy, K. Lo, A. Cohan, SciBERT: A pretrained language model for scientific text, in: K. Inui, J. Jiang, V. Ng, X. Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Association for Computational Linguistics, Hong Kong, China, 2019, pp. 3615–3620. URL: <https://aclanthology.org/D19-1371>. doi:10.18653/v1/D19-1371.
- [37] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C.H. So, J. Kang, BioBERT: a pre-trained biomedical language representation model for biomedical text mining, *Bioinformatics* 36 (2020) 1234–1240.
- [38] Y.-F. Luo, W. Sun, A. Rumshisky, MCN: a comprehensive corpus for medical concept normalization, *J. Biomed. Inform.* 92 (2019) 103132.
- [39] S. Mohan, D. Li, MedMentions: A large biomedical corpus annotated with UMLS concepts, in: *Automated Knowledge Base Construction (AKBC)*, 2018.
- [40] R. Islamaj, R. Leaman, S. Kim, D. Kwon, C.-H. Wei, D.C. Comeau, Y. Peng, D. Cisse, C. Cross, C. Fisher, et al., NLM-Chem, a new resource for chemical entity recognition in PubMed full text literature, *Sci. Data* 8 (2021) 91.
- [41] E. Voorhees, T. Alam, S. Bedrick, D. Demner-Fushman, W.R. Hersh, K. Lo, K. Roberts, I. Soboroff, L.L. Wang, TREC-COVID: constructing a pandemic information retrieval test collection, in: *ACM SIGIR Forum*, volume 54, ACM New York, NY, USA, 2021, pp. 1–12.
- [42] K. Degtyarenko, P. De Matos, M. Ennis, J. Hastings, M. Zbinden, A. McNaught, R. Alcaantara, M. Darsow, M. Guedj, M. Ashburner, ChEBI: a database and ontology for chemical entities of biological interest, *Nucleic Acids Res.* 36 (2007) D344–D350.
- [43] S. Federhen, The NCBI taxonomy database, *Nucleic Acids Res.* 40 (2012) D136–D143.
- [44] D. Maglott, J. Ostell, K.D. Pruitt, T. Tatusova, Entrez Gene: gene-centered information at NCBI, *Nucleic Acids Res.* 33 (2005) D54–D58.
- [45] Neumann, D. King, I. Beltagy, W. Ammar, ScispaCy: Fast and robust models for biomedical natural language processing. In: D. Demner-Fushman, K.B. Cohen, S. Ananiadou, J. Tsujii (Eds.), *Proceedings of the 18th BioNLP Workshop and Shared Task, Association for Computational Linguistics, Florence, Italy*, 2019, pp. 319–327. URL: <https://aclanthology.org/W19-5034>. doi:10.18653/v1/W19-5034.
- [46] S. Ravfogel, H. Taub-Tabib, Y. Goldberg. Neural Extractive Search. In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations, Association for Computational Linguistics, Online*; 2021. p. 210–217. URL: <https://aclanthology.org/2021.acl-demo.25/>.
- [47] M.A. Tanenblatt, A. Coden, I.L. Sominsky, The ConceptMapper approach to named entity recognition, *LREC* (2010) 546–551.
- [48] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert., Pre-training of Deep Bidirectional Transformers for Language Understanding, in: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (long and Short Papers)*, 2019, pp. 4171–4186.
- [49] J. Bard, S.Y. Rhee, M. Ashburner, An ontology for cell types, *Genome Biol.* 6 (2005) 1–5.
- [50] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A. P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, et al., Gene ontology: tool for the unification of biology, *Nat. Genet.* 25 (2000) 25–29.
- [51] D.A. Natale, C.N. Arighi, J.A. Blake, J. Bona, C. Chen, S.-C. Chen, K.R. Christie, J. Cowart, P. D'Eustachio, A.D. Diehl, et al., Protein Ontology (PRO): enhancing and scaling up the representation of protein entities, *Nucleic Acids Res.* 45 (2017) D339–D346.
- [52] C.J. Mungall, C. Batchelor, K. Eilbeck, Evolution of the sequence ontology terms and relationships, *J. Biomed. Inform.* 44 (2011) 87–93.
- [53] C.J. Mungall, C. Torniai, G.V. Gkoutos, S.E. Lewis, M.A. Haendel, Uberon, an integrative multi-species anatomy ontology, *Genome Biol.* 13 (2012) 1–20.
- [54] T.J. Callahan, L.J. Tripodi, L.E. Hunter, W.A. Baumgartner Jr, A framework for automated construction of heterogeneous large-scale biomedical knowledge graphs, *BioRxiv* (2020) 2020–12004.
- [55] B.T. Sherman, M. Hao, J. Qiu, X. Jiao, M.W. Baseler, H.C. Lane, T. Imamichi, W. Chang, DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update), *Nucl. Acids Res.* 50 (2022) W216–W221.
- [56] L.M. Schriml, E. Mitraka, J. Munro, B. Tauber, M. Schor, L. Nickle, V. Felix, L. Jeng, C. Bearer, R. Lichenstein, H.D. Ontology, et al., update: classification, content and workflow expansion, *Nucleic Acids Res.* 47 (2019) (2018) D955–D962.
- [57] C. Rosse, J.L. Mejino Jr, The foundational model of anatomy ontology, in: *Anatomy Ontologies for Bioinformatics: Principles and Practice*, Springer, 2008, pp. 59–117.
- [58] Disorders cluster, <https://biportal.bioontology.org/ontologies/APADISORDERS>, Accessed: 2023-11-11.
- [59] M. Zhao, J. Havrilla, J. Peng, M. Drye, M. Fecher, W. Guthrie, B. Tunc, R. Schultz, K. Wang, Y. Zhou, Development of a phenotype ontology for autism spectrum disorder by natural language processing on electronic health records, *J. Neurodev. Disord.* 14 (2022) 32.
- [60] D.M. Dooley, E.J. Griffiths, G.S. Gosal, P.L. Buttigieg, R. Hoehndorf, M.C. Lange, L.M. Schriml, F.S. Brinkman, W.W. Hsiao, FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration, *NPJ Sci. Food* 2 (2018) 23.
- [61] N.T. Nguyen, M. Miwa, Y. Tsuruoka, T. Chikayama, S. Tojo, Wide-coverage relation extraction from MEDLINE using deep syntax, *BMC Bioinf.* 16 (2015) 1–11.
- [62] UniProt Consortium, UniProt: a hub for protein information, *Nucleic Acids Res.* 43 (2015) D204–D212.
- [63] W.A. Kibbe, C. Arze, V. Felix, E. Mitraka, E. Bolton, G. Fu, C.J. Mungall, J. X. Binder, J. Malone, D. Vasant, D. Ontology, et al., update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data, *Nucleic Acids Res.* 43 (2015) (2015) D1071–D1078.
- [64] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, K. Morishima, KEGG: new perspectives on genomes, pathways, diseases and drugs, *Nucleic Acids Res.* 45 (2017) D353–D361.
- [65] C.F. Thorn, T.E. Klein, R.B. Altman, PharmGKB: the pharmacogenomics knowledge base, *Pharmacogen.: Methods Protocols* (2013) 311–320.

- [66] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T.Sajed, D. Johnson, C. Li, Z. Sayeeda, et al., DrugBank 5.0: a major update to the DrugBank database for 2018, *Nucl. Acids Res.* 46 (2018) D1074–D1082.
- [67] J. Golbeck, G. Frago, F. Hartel, J. Hendler, J. Oberthaler, B. Parsia, The National Cancer Institute's thesaurus and ontology, *J. Web Semant First Look* 1\_1\_4 (2003).
- [68] R. Leaman, C.-H. Wei, Z. Lu, tmChem: a high performance approach for chemical named entity recognition and normalization, *J. Cheminf.* 7 (2015) 1–10.
- [69] M. Gerner, G. Nenadic, C.M. Bergman, LINNAEUS: a species name identification system for biomedical literature, *BMC Bioinf.* 11 (2010) 1–17.
- [70] R. Leaman, R. Islamaj Dogan, Z. Lu, DNorm: disease name normalization with pairwise learning to rank, *Bioinformatics* 29 (2013) 2909–2917.
- [71] B. Settles, ABNER: an open source tool for automatically tagging genes, proteins and other entity names in text, *Bioinformatics* 21 (2005) 3191–3192.
- [72] D.M. Jessop, S.E. Adams, E.L. Willighagen, L. Hawizy, P. Murray- Rust, OSCAR4: a flexible architecture for chemical text-mining, *J. Cheminf.* 3 (2011) 1–12.
- [73] C.-H. Wei, H.-Y. Kao, Z. Lu, et al., GNormPlus: an integrative approach for tagging genes, gene families, and protein domains, *Biomed. Res. Int.* 2015 (2015).
- [74] C.J. Mungall, J.A. McMurry, S. Köhler, J.P. Balhoff, C. Borromeo, M. Brush, S. Carbon, T. Conlin, N. Dunn, M. Engelstad, et al., The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species, *Nucleic Acids Res.* 45 (2017) D712–D722.
- [75] S. Köhler, L. Carmody, N. Vasilevsky, J.O.B. Jacobsen, D. Danis, J.-P. Gourdine, M. Gargano, N.L. Harris, N. Matentzoglou, J.A. McMurry, et al., Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources, *Nucleic Acids Res.* 47 (2019) D1018–D1027.
- [76] V. Petri, P. Jayaraman, M. Tutaj, G.T. Hayman, J.R. Smith, J. De Pons, S. J. Laulederkind, T.F. Lowry, R. Nigam, S.-J. Wang, et al., The pathway ontology—updates and applications, *J. Biomed. Semant.* 5 (2014) 1–12.
- [77] S. Sarntivijai, Y. Lin, Z. Xiang, T.F. Meehan, A.D. Diehl, U.D. Vempati, S. C. Shiürer, C. Pang, J. Malone, H. Parkinson, et al., CLO: the cell line ontology, *J. Biomed. Semant.* 5 (2014) 1–10.
- [78] V. Sanh, L. Debut, J. Chaumont, T. Wolf, DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, *arXiv preprint arXiv:1910.01108* (2019).
- [79] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, Roberta: A robustly optimized BERT pretraining approach, *arXiv preprint arXiv:1907.11692* (2019).
- [80] E. Alsentzer, J. Murphy, W. Boag, W.-H. Weng, D. Jindi, T. Naumann, M. McDermott, Publicly available clinical BERT embeddings, in: *In: Proceedings of the 2nd Clinical Natural Language Processing Workshop*, 2019, pp. 72–78.
- [81] L.M. Schriml, C. Arze, S. Nadendla, Y.-W.-W. Chang, M. Mazaitis, V. Felix, G. Feng, W.A. Kibbe, Disease ontology: a backbone for disease semantic integration, *Nucl. Acids Res.* 40 (2012) D940–D946.
- [82] International Statistical Classification of Diseases and Related Health Problems (ICD), <https://www.who.int/tools/atc-ddd-toolkit/atc-classification>, Accessed: 2023-11-11.
- [83] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, et al., PubChem in 2021: new data content and improved web interfaces *Nucleic Acids Res.* 49 (2021) D1388–D1395.
- [84] Anatomical Therapeutic Chemical (ATC) Classification, <https://www.who.int/standards/classifications/classification-of-diseases>, Accessed: 2023-11-11.
- [85] Y. He, Y. Liu, B. Zhao, Ogg., in: *a Biological Ontology for Representing Genes and Genomes in Specific Organisms*, 2014, pp. 13–20.
- [86] C. Badenes-Olmedo, A. Alonso, O. Corcho, An overview of drugs, diseases, genes and proteins in the COR-19 corpus, *Procesamiento Del Lenguaje Natural* 69 (2022) 165–176.
- [87] I. Beltagy, K. Lo, A. Cohan, Scibert: A pretrained language model for scientific text, in: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019, pp. 3615–3620.
- [88] C. Ruiz, M. Zitnik, J. Leskovec, Identification of disease treatment mechanisms through the multiscale interactome, *Nat. Commun.* 12 (2021) 1796.
- [89] G. Michalopoulos, Y. Wang, H. Kaka, H. Chen, A. Wong, UmlsBERT: Clinical Domain Knowledge Augmentation of Contextual Embeddings Using the Unified Medical Language System Metathesaurus, in: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2021, pp. 1744–1753.
- [90] Y. Peng, Q. Chen, Z. Lu, An Empirical Study of Multi-Task Learning on BERT for Biomedical Text Mining, in: *Proceedings of the 19th SIGBioMed Workshop on Biomedical Language Processing*, 2020, pp. 205–214.
- [91] A.R. Aronson, F.-M. Lang, An overview of MetaMap: historical perspective and recent advances, *J. Am. Med. Inform. Assoc.* 17 (2010) 229–236.
- [92] T. Brown, B. Mann, N. Ryder, M. Subbiah, J.D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, et al., Language models are few-shot learners, *Adv. Neural Inf. Process. Syst.* 33 (2020) 1877–1901.
- [93] A. Chowdhery, S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H. W. Chung, C. Sutton, S. Gehrmann, et al., PALM: Scaling language modeling with pathways, *arXiv preprint arXiv:2204.02311* (2022).
- [94] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, et al., LLaMA: Open and efficient foundation language models, *arXiv preprint arXiv:2302.13971* (2023).
- [95] S. Tian, Q. Jin, L. Yeganova, P.-T. Lai, Q. Zhu, X. Chen, Y. Yang, Q. Chen, W. Kim, D. C. Comeau, et al., Opportunities and Challenges for ChatGPT and Large Language Models in Biomedicine and Health, *arXiv preprint arXiv:2306.10070* (2023).
- [96] Q. Chen, J. Du, Y. Hu, V. K. Keloth, X. Peng, K. Raja, R. Zhang, Z. Lu, H. Xu, Large language models in biomedical natural language processing: benchmarks, baselines, and recommendations, *arXiv preprint arXiv:2305.16326* (2023).
- [97] B. Jimenez Gutierrez, N. McNeal, C. Washington, Y. Chen, L. Li, H. Sun, Y. Su, Thinking about GPT-3 in-context learning for biomedical IE? think again, in: Y. Goldberg, Z. Kozareva, Y. Zhang (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2022*, Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, 2022, pp. 4497–4512. URL: <https://aclanthology.org/2022.findings-emnlp.329>. doi:10.18653/v1/2022.findings-emnlp.329.
- [98] M. Agrawal, S. Hegselmann, H. Lang, Y. Kim, D. Sontag, Large language models are few-shot clinical information extractors, in: Y. Goldberg, Z. Kozareva, Y. Zhang (Eds.), *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, 2022, pp. 1998–2022. URL: <https://aclanthology.org/2022.emnlp-main.130>. doi:10.18653/v1/2022.emnlp-main.130.
- [99] H. W. Chung, L. Hou, S. Longpre, B. Zoph, Y. Tay, W. Fedus, Y. Li, X. Wang, M. Dehghani, S. Brahma, et al., Scaling instruction-finetuned language models, *arXiv preprint arXiv:2210.11416* (2022).
- [100] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q.V. Le, D. Zhou, et al., Chain-of-thought prompting elicits reasoning in large language models, *Adv. Neural Inf. Process. Syst.* 35 (2022) 24824–24837.
- [101] S. Wadhwa, S. Amir, B. Wallace, Revisiting relation extraction in the era of large language models, in: A. Rogers, J. Boyd-Graber, N. Okazaki (Eds.), *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Association for Computational Linguistics, Toronto, Canada, 2023, pp. 15566–15589. URL: <https://aclanthology.org/2023.acl-long.868>. doi:10.18653/v1/2023.acl-long.86.
- [102] Wang Q, Downey D, Ji H, Hope T. Learning to Generate Novel Scientific Directions with Contextualized Literature-based Discovery. *arXiv preprint arXiv:2305.14259*. 2023 May 23.
- [103] Zhu Y, Wang X, Chen J, Qiao S, Ou Y, Yao Y, Deng S, Chen H, Zhang N. LLMs for Knowledge Graph Construction and Reasoning: Recent Capabilities and Future Opportunities. *arXiv preprint arXiv:2305.13168*. 2023 May 22.
- [104] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.T. Yih, T. Rocktäschel, S. Riedel, Retrieval-augmented generation for knowledge-intensive NLP tasks, *Adv. Neural Inform. Process. Syst.* 33 (2020) 9459–9474.
- [105] Soman K, Rose PW, Morris JH, Akbas RE, Smith B, Peetoom B, Villouta-Reyes C, Ceroni G, Shi Y, Rizk-Jackson A, Israni S, et al. Biomedical knowledge graph-enhanced prompt generation for large language models. *arXiv preprint arXiv:2311.17330*. 2023 Nov 29.
- [106] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y.J. Bang, A. Madotto, P. Fung, Survey of hallucination in natural language generation, *ACM Comput. Surv.* 55 (2023) 1–38.
- [107] P. Hitzler, A. Eberhart, M. Ebrahimi, M. K. Sarker, L. Zhou, Neurosymbolic approaches in artificial intelligence, *Natl. Sci. Rev.* 9 (2022) nwac035.
- [108] S. Pan, L. Luo, Y. Wang, C. Chen, J. Wang, X. Wu, Unifying large language models and knowledge graphs: a roadmap, *arXiv preprint arXiv:2306.08302* (2023).

Halil Kilicoglu<sup>a,\*</sup>, Faezeh Ensan<sup>b</sup>, Bridget McInnes<sup>c</sup>, Lucy Lu Wang<sup>d</sup>

<sup>a</sup> School of Information Sciences, University of Illinois Urbana Champaign, Champaign, IL, USA

<sup>b</sup> Department of Electrical, Computer, and Biomedical Engineering, Toronto Metropolitan University, Toronto, ON, Canada

<sup>c</sup> Department of Computer Science, Virginia Commonwealth University, Richmond, VA, USA

<sup>d</sup> Information School, University of Washington, Seattle, WA, USA

\* Corresponding author at: 501 E Daniel Street, Champaign, IL, 61820, USA.

E-mail addresses: [halil@illinois.edu](mailto:halil@illinois.edu) (H. Kilicoglu), [fensan@torontomu.ca](mailto:fensan@torontomu.ca) (F. Ensan), [btmcinnes@vcu.edu](mailto:btmcinnes@vcu.edu) (B. McInnes), [lucylw@uw.edu](mailto:lucylw@uw.edu) (L.L. Wang).