Web Semantics: Science, Services and Agents on the World Wide Web 84 (2025) 100851

Contents lists available at ScienceDirect

Web Semantics: Science, Services and Agents on the World Wide Web

journal homepage: www.elsevier.com/locate/websem

Towards leveraging explicit negative statements in knowledge graph embeddings

Rita T. Sousa [a,*], Catia Pesquita [b], Heiko Paulheim [a]

[a] Data and Web Science Group, University of Mannheim, Mannheim, Germany
[b] LASIGE, University of Lisbon, Lisbon, Portugal

ARTICLE INFO

Keywords:
Knowledge graph
Knowledge graph embedding
Negative statements

ABSTRACT

Knowledge Graphs are used in various domains to represent knowledge about entities and their relations. In the vast majority of cases, they capture what is known to be true about those entities, i.e., positive statements, while the Open World Assumption implicitly states that everything not expressed in the graph may or may not be true. This makes it difficult and less frequent to capture information explicitly known not to be true, i.e., negative statements. Moreover, while those negative statements could bear the potential to learn more useful representations in knowledge graph embeddings, that direction has been explored only rarely. However, in many domains, negative information is particularly interesting, for example, in recommender systems, where negative associations of users and items can help in learning better user representations, or in the biomedical domain, where the knowledge that a patient does exhibit a specific symptom can be crucial for accurate disease diagnosis.

In this paper, we argue that negative statements should be given more attention in knowledge graph embeddings. Moreover, we investigate how they can be used in knowledge graph embedding methods, highlighting their potential in some interesting use cases. We discuss some existing works and preliminary results that incorporate explicitly declared negative statements in walk-based knowledge graph embedding methods. Finally, we outline promising avenues for future research in this area.

1. Introduction

Knowledge Graphs (KGs) capture information about real-world entities and their relationships in a structured format that is understandable by both machines and humans [1]. Over the years, several KGs have been developed, both freely accessible and commercially available. Popular examples include the Google Knowledge Graph [2], DBpedia [3], and Wikidata [4]. As KGs have gained popularity, various KG embedding methods have emerged and have been successfully applied across numerous fields [5]. These methods project KGs into low-dimensional spaces while preserving the KG's structural or semantic characteristics, allowing embeddings to serve as features in machine learning tasks or to assess semantic similarity between entities [5–7]. The impact of KG embeddings is growing alongside the increasing volume and complexity of data in KGs.

KGs predominantly represent facts as positive statements across diverse domains. For example, a fact might express that Albert Einstein was awarded a Nobel Prize in Physics. While less common, KGs can also represent facts as negative statements. In [8], two types of negative statements, *grounded* and *universal*, are distinguished. Grounded negative statements assert that a specific relationship does not exist between two entities, such as stating that Stephen Hawking was not awarded the Nobel Prize in Physics, expressed as $\neg(\text{Stephen Hawking}, \text{awarded}, \text{Nobel Prize in Physics})$. Therefore, a grounded negative statement $\neg(s,p,o)$ is satisfied if $(s,p,o) \notin KG$. In OWL [9], such statements are represented using a class restriction applied to a single individual, or, since OWL2, using negative object property assertions, which state that an entity is not connected by a specific object property expression to another entity [10]. In contrast, universally negative statements convey negation at the level of a specific subject and predicate, indicating that a particular entity does not have any relationship of that type. For instance, stating that Isaac Newton was never married, expressed as $\neg\exists o : (\text{Isaac Newton}, \text{married}, o)$. A universally negative statement $\neg\exists o : (s,p,o)$ is satisfied if there exists no o such that $(s,p,o) \in KG$.

Most real-world KGs operate under the Open World Assumption, where non-stated facts may represent missing/unknown facts or true negative statements. In contrast, many AI systems are designed to operate under the Closed World assumption or, in certain cases, under the Local Closed World assumption [11], also known as the Partial Completeness assumption [12]. According to the Local Closed World

* Corresponding author.
E-mail address: rita.sousa@uni-mannheim.de (R.T. Sousa).

https://doi.org/10.1016/j.websem.2024.100851
Received 20 November 2024; Received in revised form 7 December 2024; Accepted 17 December 2024
Available online 24 December 2024

 Journal of Web Semantics 84 (2025) 100851

assumption, if there exists a triple (s, p, o), then the KG is assumed to contain all objects for any statements of the form $(s, p, ?)$. Although this is well-suited for relationships where there is only one object (e.g., birthIn or diedIn), this assumption can break down when the relationships associate multiple objects with each subject (e.g., hasFriend or worksAt), which is common for most relationships [12]. This dichotomy creates challenges when integrating AI systems with KGs as assumptions diverge.

Consequently, explicitly declared negative statements are then increasingly recognized as valuable in KGs, with a growing recognition of their importance for knowledge representation [13]. Several approaches have been developed to enrich KGs by incorporating interesting and meaningful negative statements. Since the set of correct negative statements is near-infinite, the interesting and meaningful negative statements are those that are expected or believed to be true but later turn out to be false. For instance, within the context of the Nobel Prize, one could theoretically add negative statements for all individuals who lack a positive statement about winning a Nobel Prize, regardless of their field or background. However, such negative statements become more interesting and meaningful when they challenge expectations. For example, the negative statement stating that Stephen Hawking was not awarded the Nobel Prize in Physics contrasts with the widespread expectation that this renowned theoretical physicist would have received such recognition. These approaches have been applied not only to general-purpose KGs like Wikidata [14], but also in the biomedical context [15,16].

Despite the advancements in enriching KGs with meaningful negative statements, the importance of exploring this type of information in KG embedding methods is still in the early stages. Most existing embedding methods do not explore and integrate explicitly declared negative statements into their frameworks, often overlooking the underlying ontological implications. This gap in handling negative information leads to less accurate entity representations, potentially limiting the embeddings' effectiveness in downstream applications. To address these limitations, future research must prioritize the exploration of negative statements into embedding techniques, potentially leading to richer and more robust representations.

In this paper, we argue that negative statements can strengthen KG embedding methods and identify several key use cases that reinforce the value of such statements in KGs. We also conduct a comprehensive analysis of existing approaches. Through this analysis, we identify limitations in current methods, and we propose actionable directions for future research.

2. Embedding spaces with negative statements

KG embedding models aim to represent entities and relationships as low-dimensional vectors that capture the semantics and relationships within the KG. Most KG embedding methods involve generating a set of triples that are expected to have a low probability of truth and using these as negative examples to model training. The Local Closed World assumption is inherently applied in the generation of negative samples, when modifying the object in a statement (s, p, o) to create (s, p, o'), the new triple is considered a negative example if it does not exist in the KG. In an ideal scenario, these latent representations would encode meaningful semantics: entities that share similar attributes or relationships are positioned close together in the vector space, while entities with distinct attributes or relationships are located farther apart. For example, entities like *Stephen Hawking* and *Albert Einstein* and *J. Robert Oppenheimer* should have embeddings close together in vector space because they share similar attributes, such as being prominent scientists. In contrast, an entity like *James Stewart*, an *actor*, should have an embedding positioned farther away from these scientists.

In a KG that contains both positive and meaningful negative statements, effective KG embedding models also need to capture the negative semantics. Entities with similar negative statements (e.g., entities

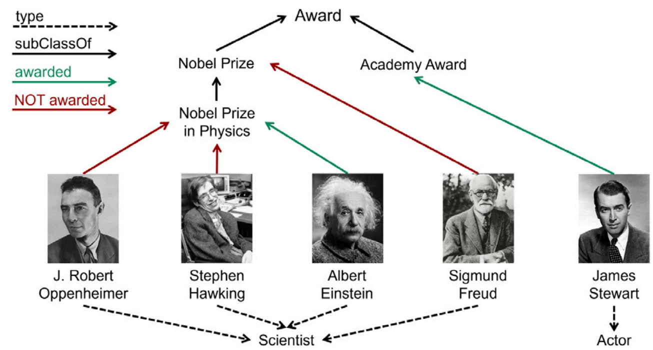

Fig. 1. A simple example motivating the negative statements challenge.

known for lacking certain attributes) should also be close in the vector space. Moreover, if two pairs of entities share the same positive attributes, but the second pair also has contradictory attributes, they should be farther apart than the first pair. Using the previous example, both *Stephen Hawking* and *J. Robert Oppenheimer* surprisingly never received a *Nobel Prize in Physics*, a shared negative statement. In contrast, *Albert Einstein* was awarded the *Nobel Prize in Physics* in 1921. This should be reflected in their embeddings: *Stephen Hawking* and *J. Robert Oppenheimer* should be closer to each other and more distant from *Albert Einstein* in the embedding space, capturing their shared attribute of non-award status and distinguishing them from Einstein's positive statement in this context. Fig. 1 illustrates this example.

Additionally, when working with ontology-rich KGs, additional challenges arise. In such ontology-rich KGs, very common in the biomedical domain, an ontology is used to provide rich descriptions of real-world entities rather than focusing on the relationships between entities themselves, resulting in a very rich TBox, with a comparatively simpler ABox. When real-world entities in these KGs are described through both positive and negative statements, a key distinction between these two types of statements lies in the inheritance of properties from the superclasses or subclasses associated with the assigned ontology class. For example, considering an ontology describing awards, a hierarchical class structure might define the *Nobel Prize in Physics* as a subclass of *Nobel Prize*, which in turn is a subclass of *Award*. A positive statement asserting that *Albert Einstein* received the *Nobel Prize in Physics* implies that he also received an *Award*, as class assignments propagate up through the superclass hierarchy. Conversely, a negative statement indicating that *Sigmund Freud* did not receive the *Nobel Prize* - despite being nominated for both the Nobel Prize in Physiology or Medicine and the Nobel Prize in Literature - would imply that he did not receive the *Nobel Prize in Physics* or any other Nobel Prize in a different field, as the negation would propagate down the subclass hierarchy. This cascading effect of negation means that a negative assertion about a class also applies to all its subclasses, while a positive assertion about a class applies to all its superclasses.

Since OWL ontologies primarily define taxonomies through subclass relationships, typical KG embedding methods struggle to capture reverse paths that would better capture negative assertions. For instance, in path-based embedding methods like RDF2vec [17], the path *Albert Einstein → awarded → Nobel Prize in Physics → subClassOf → Nobel Prize → subClassOf → Award* and the path *Stephen Hawking → not awarded → Nobel Prize in Physics → subClassOf → Nobel Prize → subClassOf → Award* appear almost identical. Without a robust way to differentiate between the *awarded* and *not awarded* properties, these paths may lead to similar embeddings, undermining the KG's ability to represent negative information accurately. To address these challenges, KG embedding methods that can recognize and properly encode the semantics of negative statements are essential.

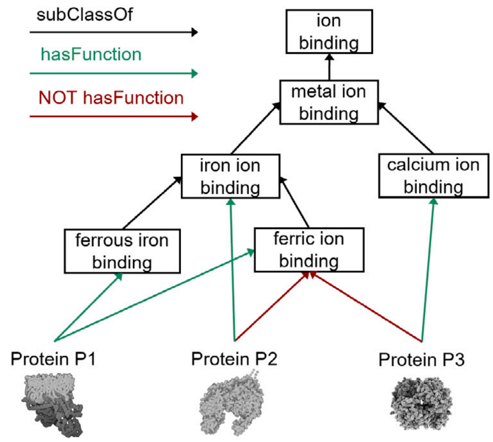

Fig. 2. Subgraph of a biomedical KG that represents proteins on the functions they perform and do not perform. In this example, three proteins (P1, P2, P3) are represented. Proteins P1 and P2 demonstrate high semantic similarity as they share both positive functional annotations (performing metal ion binding) and negative annotations (not performing ferric ion binding). This dual perspective enhances the understanding of their biological roles and highlights a higher likelihood of functional interaction.

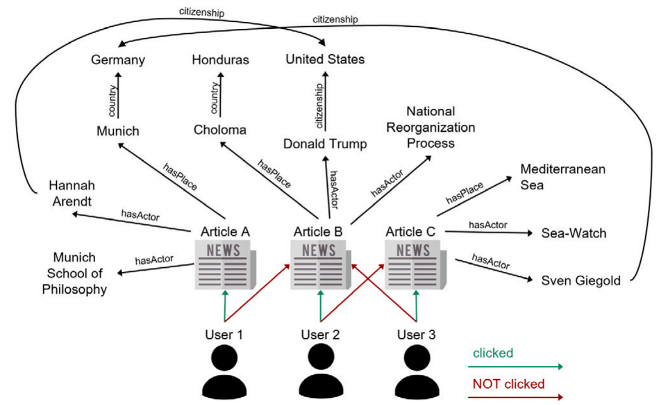

Fig. 3. Subgraph containing associations between users and newspaper articles, incorporating both positive (articles clicked) and negative (articles not clicked) associations. By examining these two types of associations for each user, we can gain insights into their preferences. For instance, User 1 and User 3 exhibit similar taste profiles. Based on this, we can recommend Article C, which User 3 positively engaged with, to User 1.

3. Use cases and benchmarks

The inability to differentiate between a statement being false or unknown creates significant challenges for knowledge representation across different fields. In this section, we focus particularly on three use cases involving biomedical KGs [18], KGs for news recommendation [19], and food KGs [20], where incorporating the negative statements into embedding generation could have a significant impact.

In biomedical KGs, entities such as genes, proteins, and diseases are characterized through statements linking them to ontology classes, often referred to as annotations. A well-known example is the Gene Ontology (GO) KG [21], where a positive statement can denote that a gene product performs a function as described by the GO, while a negative statement can indicate that a gene product does not perform a specific function that is typically performed by its homologs [15] (Fig. 2). Some studies have already shown the importance of considering negative statements in GO KG. For instance, [15] showed that a balance between positive and negative annotations for proteins results in a more robust evaluation of protein function prediction using orthology-based methods. Similarly, [16] reported that the incorporation of negative statements enhances the prediction performance of protein functions using a Gaussian random field-based label propagation algorithm. However, most approaches that generate KG embeddings and then use them to train machine learning models to predict associations between genes and diseases or interactions between proteins [7,22–24] rely solely on positive statements from GO KG. Recent work by [18] introduced a collection of datasets for various biomedical tasks – such as protein-protein interaction prediction, disease prediction, and gene-disease association prediction – that include KGs with both positive and negative statements. Their study compared the impact of using KG with only positive statements versus both positive and negative statements on the proposed datasets, employing two popular KG embedding methods. The results revealed that the added information given by negative statements does not always improve the performance of biomedical tasks. These findings suggest that, while negative information is useful, embedding methods must be designed to effectively capture both what proteins and genes can and what they cannot. Such methods should capture a more accurate measure of similarity, potentially leading to improved prediction performance.

Beyond biomedical applications, another compelling use case of negative statements is recommender systems, an area where KG embedding methods have attracted considerable attention [25–27]. When based on a KG, the edges connecting a user to an item can be positive (user viewed/liked an item) or negative (user viewed/did not like an item), as shown in Fig. 3. Besides the positive clicks, unclicked articles may offer critical information about content that users actively avoid, potentially indicating deeper biases or political leanings. Therefore, user embeddings should accurately reflect both positive and negative associations between a user and an item. One example dataset containing such positive and negative edges is the NeMig KG [19]. The NeMig KG integrates data from news articles, metadata, linked Wikidata entities, and user interaction data collected through online studies. These studies captured user demographics, political views, and explicit click feedback on a selection of articles. While previous experiments on the dataset have shown that recommendations on the KG can lead to very good results, and, at the same time, mitigate biases in pure text-based news recommendation [28], the algorithm utilized in the paper, i.e., RippleNet [29], like most recommender algorithms, does not distinguish between unrated and negatively rated items. Here, representation learning algorithms can learn more expressive user models.

A third example use case pertains to food KGs. As poor dietary habits are linked to various health conditions, there has been growing interest in structuring knowledge about food and its ingredients. Consequently, several ontologies and KGs [20,30,31] have been proposed, and KG embedding methods are being applied to tasks such as food recommendation or ingredient substitution [32–34]. FoodKG [20] exemplifies a KG that provides information about recipes, their ingredients, and the nutritional content of individual food items. Such KGs can be further enriched with negative statements, such as details about ingredients that are explicitly excluded from a recipe or are not associated with a particular food item (Fig. 4). Incorporating this type of information can be particularly valuable for improving recommendations or suggesting substitute ingredients according to food allergy restrictions or dietary preferences, such as vegan or gluten-free diets. However, all the works discussed above for recipe recommendation and ingredient substitution use algorithms incapable of utilizing such negative statements. With those algorithms, the recommendations will always underperform when taking negative information into account. For example, with a standard KG embedding algorithm, two vegan recipes with dissimilar ingredients will always be far away in the vector space, while two recipes with similar ingredients, where one is vegan, and one is not (e.g., vegan curry and chicken curry) will be close in the vector space. Depending on the use case at hand, this is an undesirable property. In particular, users following a vegan diet will be unsatisfied with such a system.

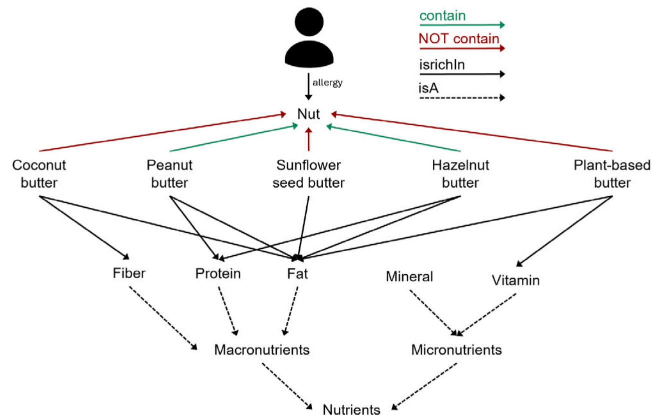

Fig. 4. Subgraph within the food domain where different types of butter are described based on nutrients richness and whether or not they contain nuts. These negative statements are particularly useful in scenarios like identifying suitable butter substitutes for individuals with nut allergies. By incorporating the negative statements into embedding generation, nut-based butter can be effectively distinguished from nut-free alternatives in a latent vector space, enabling models to prioritize and recommend substitutes that align with specific dietary constraints.

4. Existing approaches and limitations

Numerous KG embedding models have been proposed in recent years [5,35]. Two main research directions have emerged in this field [36].

The first focuses on link prediction, by learning embeddings that assign lower scores to valid triples compared to invalid ones. In these approaches, the embedding model is able to distinguish correct triples from incorrect ones by optimizing for a scoring function that reflects the plausibility of each triple. KG embedding methods belonging to the first research direction primarily differ in three core components: representation space, encoding model, and loss function. These differences lead to different families of approaches: translational distance models [37–40], geometric models [23,41], semantic matching models [42–46], and deep learning-based models [47,48].

The second direction focuses on embedding entities in the KG for use in downstream tasks. In this second direction, walk-based methods are especially prominent. The majority of these approaches rely on generating entity sequences by performing walks on the graph, creating a corpus of sequences analogous to a collection of word sentences. Then, these sequences are fed into a neural language model, which learns a latent low-dimensional representation for each entity within the sequence corpus. The second research direction includes embedding methods like RDF2Vec [17], OWL2Vec* [49] and OPA2Vec [24]. More recently, a novel approach called TrueWalks [50] was introduced to incorporate negative statements into the KG representation learning process.

TrueWalks is, to the best of our knowledge, the only approach to consider a graph that contains explicitly declared negative statements and their semantic implications. This is fundamentally different from other KG embedding methods that produce negative examples by negative random sampling strategies to train representations based on positive examples that bring the representations of nodes that are linked closer while distancing them from the negative examples. While these methods can handle negative statements in a KG, they treat them as any other type of statement. Additionally, TrueWalks is distinct from approaches that consider weighted KGs [46,51–54] where each triple is associated with a confidence score that reflects its plausibility. While weighted KGs account for uncertainty, the lower end of their range of weights (usually 0) reflects minimum plausibility, which is equivalent to a non-existing statement in a KG under open world semantics. However, this is different from those meaningful negative statements, which explicitly refer to triples that are known to be false.

TrueWalks proposes a novel walk-generation method that distinguishes between positive and negative statements while also capturing the semantic implications of negation, especially in ontology-rich KGs. It generates biased walks in the KG: a positive statement implies that whenever a subclass edge is found, it is traversed from subclass to superclass, whereas a negative statement results in a traversal of subclass edges in the opposite direction. This way, this approach effectively splits walks into two sets, one for positive and another for negative statements, enabling the model to learn dual latent representations, separately capturing the positive and the negative aspects. TrueWalks was extensively evaluated on established benchmarks and compared with several KG embedding methods across two different experiments. In the first, KG embedding methods were applied to a KG containing only positive statements. In the second, a KG with both positive and negative statements was used, with negative statements declared as an object property so that embedding methods distinguish the two types of statements as two distinct types of relation. The results showed that incorporating the negative statements, even as any other type of statement, generally improved the performance of most KG embedding methods. However, TrueWalks outperformed all others, likely due to its unique capability to account for the semantic implications of inheritance. TrueWalks still has some limitations that need to be addressed. For instance, it is primarily effective for KGs supported by rich ontologies, as the biased walks rely on the assumption that entities are linked to ontology classes that have well-defined subclass and superclass relationships. Additionally, TrueWalks does not account for how contradictory statements might affect the dissimilarity between entities since it learns a dual representation.

While not directly tackling the negative statement statements directly, semantic information has been progressively incorporated into several KG embedding methods over recent years, particularly through modifications in the loss function [40,55–57] or through regularization techniques that enforce the embedding space [58,59]. All of these existing approaches showed promising results in embedding semantic information effectively. Therefore, this represents a fertile area for future research, as these techniques could potentially be adapted to dealing with negative property assertions or manipulate the embedding space to separate entities with contradictory statements.

Another promising area of research focuses on enhancing negative sampling techniques, which play a critical role in the training of KG embedding methods [60,61]. Negative sampling involves generating a set of triples that are expected to have a low probability of truth and using these as negative examples. The meaningful negative statements contrast with these negative examples that are essential for effective model training. However, as demonstrated in [61], exploring semantically valid and invalid negative examples can be beneficial and treating these two types of negative examples within the loss function has shown promise. This opens further opportunities to investigate how meaningful statements can be considered as an additional type of negative example and integrated into the loss function.

5. Conclusions

KGs with negative statements have been increasingly recognized as valuable [8]. As a result, several KGs containing meaningful negative statements are now available [14], along with established benchmarks for evaluating tasks on these KGs [18]. This provides a solid foundation for further research and experimentation. Furthermore, some works have already shown the advantage of including negative statements in KG embedding-based applications [50]. By explicitly considering negative statements, we can learn more accurate and expressive entity representations, which in turn leads to better performance in tasks such as link prediction and recommendation systems. However, the use of such statements within KG embedding approaches remains largely unexplored, and there is a pressing need for novel approaches.

In this paper, we have shown three exemplary use cases from different domains where we argue that respecting negative information in KG embeddings will bring advantages. Research questions for such embedding methods include, but are not limited to:

1. Which representation mechanism of negative statements can be best exploited by KG embedding methods?
2. What characteristics do different settings of negative statements have (e.g., fraction of positive and negative statements in a KG, contradicting vs. contradiction-free statements, distribution of negative statements)?
3. Out of the exploitation strategies discussed above (e.g., adapting the loss function, changing negative sampling strategies, adapting walk strategies in walk-based embedding approaches), which ones work best in those settings?
4. How can negative statements help a reasoner in the loop of the embedding process?
5. When considering the link prediction task, can we build models that simultaneously predict positive and negative links? Can the two tasks benefit from each other?

These questions open a vast world of opportunities, whether by extending existing methods to consider negative statements, reusing some existing paradigms or proposing completely novel approaches that are designed specifically for KGs containing both positive and negative statements. Ultimately, these advancements have the potential to significantly leverage the utility of KGs in a wide range of real-world scenarios.

CRediT authorship contribution statement

Rita T. Sousa: Writing – original draft. **Catia Pesquita:** Writing – review & editing. **Heiko Paulheim:** 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.

Data availability

No data was used for the research described in the article.

References

[1] A. Hogan, E. Blomqvist, M. Cochez, C. d'Amato, G.d. Melo, C. Gutierrez, S. Kirrane, J.E.L. Gayo, R. Navigli, S. Neumaier, et al., Knowledge graphs, ACM Comput. Surv. 54 (4) (2021) 1–37.
[2] A. Singhal, et al., Introducing the knowledge graph: things, not strings, Off. Google Blog 5 (16) (2012) 3.
[3] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, Z. Ives, Dbpedia: A nucleus for a web of open data, in: International Semantic Web Conference, 2007, pp. 722–735.
[4] D. Vrandečić, M. Krötzsch, Wikidata: a free collaborative knowledgebase, Commun. ACM 57 (10) (2014) 78–85.
[5] Q. Wang, Z. Mao, B. Wang, L. Guo, Knowledge graph embedding: A survey of approaches and applications, IEEE Trans. Knowl. Data Eng. 29 (12) (2017) 2724–2743.
[6] S.K. Mohamed, A. Nounu, V. Nováček, Biological applications of knowledge graph embedding models, Brief. Bioinform. 22 (2) (2021) 1679–1693.
[7] M. Kulmanov, F.Z. Smaili, X. Gao, R. Hoehndorf, Semantic similarity and machine learning with ontologies, Brief. Bioinform. 22 (4) (2021) bbaa199.
[8] H. Arnaout, S. Razniewski, G. Weikum, J.Z. Pan, Negative statements considered useful, J. Web Semant. 71 (2021) 100661.
[9] F. Van Harmelen, D.L. McGuinness, OWL web ontology language overview, World Wide Web Consortium (W3C) Recomm. 69 (2004) 70.
[10] O. Noppens, Negative property assertion pattern (npas), in: Workshop on Ontology Patterns, 2009, p. 120.
[11] P. Doherty, W. Lukaszewicz, A. Szalas, Efficient reasoning using the local closed-world assumption, in: International Conference on Artificial Intelligence: Methodology, Systems, and Applications, 2000, pp. 49–58.
[12] L.A. Galárraga, C. Teflioudi, K. Hose, F. Suchanek, AMIE: association rule mining under incomplete evidence in ontological knowledge bases, in: International Conference on World Wide Web, 2013, pp. 413–422.
[13] G. Flouris, Z. Huang, J.Z. Pan, D. Plexousakis, H. Wache, Inconsistencies, negations and changes in ontologies, in: National Conference on Artificial Intelligence, Vol. 2, 2006, pp. 1295–1300.
[14] H. Arnaout, S. Razniewski, G. Weikum, J.Z. Pan, Negative knowledge for open-world wikidata, in: Companion Proceedings of the Web Conference 2021, 2021, pp. 544–551.
[15] A. Warwick Vesztrocy, C. Dessimoz, Benchmarking gene ontology function predictions using negative annotations, Bioinformatics 36 (Supplement_1) (2020) i210–i218.
[16] G. Fu, J. Wang, B. Yang, G. Yu, NegGOA: negative GO annotations selection using ontology structure, Bioinformatics 32 (19) (2016) 2996–3004.
[17] P. Ristoski, H. Paulheim, RDF2Vec: RDF graph embeddings for data mining, in: International Semantic Web Conference, Springer, 2016, pp. 498–514.
[18] R.T. Sousa, S. Silva, C. Pesquita, Benchmark datasets for biomedical knowledge graphs with negative statements, in: Workshop on Semantic Web Solutions for Large-Scale Biomedical Data Analytics Co-Located with Extended Semantic Web Conference, 2023.
[19] A. Iana, M. Alam, A. Grote, K. Luwig, P. Müller, C. Weinhardt, H. Paulheim, Nemig-a bilingual news collection and knowledge graph about migration, in: Workshop on News Recommendation and Analytics Co-Located with RecSys, 2023.
[20] S. Haussmann, O. Seneviratne, Y. Chen, Y. Ne'eman, J. Codella, C.-H. Chen, D.L. McGuinness, M.J. Zaki, Foodkg: a semantics-driven knowledge graph for food recommendation, in: International Semantic Web Conference, Springer, 2019, pp. 146–162.
[21] GO Consortium, The gene ontology resource: 20 years and still GOing strong, Nucleic Acids Res. 47 (D1) (2018) D330–D338.
[22] S.B. Zhang, Q.R. Tang, Protein–protein interaction inference based on semantic similarity of gene ontology terms, J. Theoret. Biol. 401 (2016) 30–37.
[23] M. Kulmanov, W. Liu-Wei, Y. Yan, R. Hoehndorf, EL embeddings: geometric construction of models for the description logic EL++, in: International Joint Conference on Artificial Intelligence, 2019.
[24] F.Z. Smaili, X. Gao, R. Hoehndorf, OPA2Vec: combining formal and informal content of biomedical ontologies to improve similarity-based prediction, Bioinformatics 35 (12) (2019) 2133–2140.
[25] J.C. Zhang, A.M. Zain, K.Q. Zhou, X. Chen, R.M. Zhang, A review of recommender systems based on knowledge graph embedding, Expert Syst. Appl. (2024) 123876.
[26] Q. Guo, F. Zhuang, C. Qin, H. Zhu, X. Xie, H. Xiong, Q. He, A survey on knowledge graph-based recommender systems, IEEE Trans. Knowl. Data Eng. 34 (8) (2020) 3549–3568.
[27] R. Sun, X. Cao, Y. Zhao, J. Wan, K. Zhou, F. Zhang, Z. Wang, K. Zheng, Multi-modal knowledge graphs for recommender systems, in: ACM International Conference on Information & Knowledge Management, 2020, pp. 1405–1414.
[28] M. Alam, A. Iana, A. Grote, K. Ludwig, P. Müller, H. Paulheim, Towards analyzing the bias of news recommender systems using sentiment and stance detection, in: Companion Proceedings of the Web Conference 2022, 2022, pp. 448–457.
[29] H. Wang, F. Zhang, J. Wang, M. Zhao, W. Li, X. Xie, M. Guo, Ripplenet: Propagating user preferences on the knowledge graph for recommender systems, in: ACM International Conference on Information and Knowledge Management, 2018, pp. 417–426.
[30] C. Fu, Z. Huang, F. van Harmelen, T. He, X. Jiang, Food4healthkg: Knowledge graphs for food recommendations based on gut microbiota and mental health, Artif. Intell. Med. 145 (2023) 102677.
[31] 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 (1) (2018) 23.
[32] S.S. Shirai, O. Seneviratne, M.E. Gordon, C.H. Chen, D.L. McGuinness, Semantics-Driven ingredient substitution in the foodKG? in: International Semantic Web Conference (Posters/Demos/Industry), CEUR-WS, 2020.
[33] J. Loesch, L. Meeckers, I. van Lier, A. de Boer, M. Dumontier, R. Celebi, Automated identification of food substitutions using knowledge graph embeddings, in: International Semantic Web Applications and Tools for Health Care and Life Sciences Conference, 2022, pp. 19–28.
[34] S.S. Shirai, O. Seneviratne, M.E. Gordon, C.H. Chen, D.L. McGuinness, Identifying ingredient substitutions using a knowledge graph of food, Front. Artif. Intell. 3 (2021) 621766.
[35] J. Cao, J. Fang, Z. Meng, S. Liang, Knowledge graph embedding: A survey from the perspective of representation spaces, ACM Comput. Surv. 56 (6) (2024) 1–42.

[36] J. Portisch, N. Heist, H. Paulheim, Knowledge graph embedding for data mining vs. knowledge graph embedding for link prediction–two sides of the same coin? Semant. Web (2022) 1–24.
[37] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, O. Yakhnenko, Translating embeddings for modeling multi-relational data, Adv. Neural Inf. Process. Syst. 26 (2013).
[38] Z. Wang, J. Zhang, J. Feng, Z. Chen, Knowledge graph embedding by translating on hyperplanes, in: AAAI Conference on Artificial Intelligence, AAAI Press, 2014, pp. 1112–1119.
[39] Y. Lin, Z. Liu, M. Sun, Y. Liu, X. Zhu, Learning entity and relation embeddings for knowledge graph completion, in: AAAI Conference on Artificial Intelligence, Vol. 29, No. 1, AAAI Press, 2015.
[40] C. d'Amato, N.F. Quatraro, N. Fanizzi, Injecting background knowledge into embedding models for predictive tasks on knowledge graphs, in: Extended Semantic Web Conference, Springer, 2021, pp. 441–457.
[41] B. Xiong, N. Potyka, T.K. Tran, M. Nayyeri, S. Staab, Faithful Embeddings for EL++ Knowledge Bases, in: International Semantic Web Conference, Springer, 2022, pp. 22–38.
[42] M. Nickel, V. Tresp, H.P. Kriegel, A three-way model for collective learning on multi-relational data, in: International Conference on International Conference on Machine Learning, 2011.
[43] B. Yang, W. tau Yih, X. He, J. Gao, L. Deng, Embedding entities and relations for learning and inference in knowledge bases, 2015.
[44] M. Nickel, L. Rosasco, T. Poggio, Holographic embeddings of knowledge graphs, in: AAAI Conference on Artificial Intelligence, AAAI Press, 2016.
[45] T. Trouillon, J. Welbl, S. Riedel, É. Gaussier, G. Bouchard, Complex Embeddings for Simple Link Prediction, in: International Conference on International Conference on Machine Learning, Vol. 48, JMLR.org, 2016, pp. 2071–2080.
[46] X. Chen, M. Chen, W. Shi, Y. Sun, C. Zaniolo, Embedding uncertain knowledge graphs, in: AAAI Conference on Artificial Intelligence, Vol. 33, No. 01, AAAI Press, 2019, pp. 3363–3370.
[47] D. Wang, P. Cui, W. Zhu, Structural deep network embedding, in: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, 2016.
[48] S. Cao, W. Lu, Q. Xu, Deep neural networks for learning graph representations, in: AAAI Conference on Artificial Intelligence, AAAI Press, 2016.
[49] J. Chen, P. Hu, E. Jimenez-Ruiz, O.M. Holter, D. Antonyrajah, I. Horrocks, OWL2Vec*: Embedding of OWL ontologies, Mach. Learn. (2021) 1–33.
[50] R.T. Sousa, S. Silva, H. Paulheim, C. Pesquita, Biomedical knowledge graph embeddings with negative statements, in: International Semantic Web Conference, Springer, 2023, pp. 428–446.
[51] X. Wu, H. Pang, Y. Fan, Y. Linghu, Y. Luo, ProbWalk: A random walk approach in weighted graph embedding, Procedia Comput. Sci. 183 (2021) 683–689.
[52] M. Nayyeri, G.M. Cil, S. Vahdati, F. Osborne, A. Kravchenko, S. Angioni, A. Salatino, D. Reforgiato Recupero, E. Motta, J. Lehmann, Link prediction of weighted triples for knowledge graph completion within the scholarly domain, IEEE Access 9 (2021) 116002–116014.
[53] K.W. Kun, X. Liu, T. Racharak, L.M. Nguyen, TransHExt: a weighted extension for transh on weighted knowledge graph embedding, in: International Semantic Web Conference (Posters/Demos/Industry), 2022.
[54] Z.M. Chen, M.Y. Yeh, T.W. Kuo, PASSLEAF: a pool-based semi-supervised learning framework for uncertain knowledge graph embedding, in: AAAI Conference on Artificial Intelligence, Vol. 35, No. 5, AAAI Press, 2021, pp. 4019–4026.
[55] P. Minervini, L. Costabello, E. Muñoz, V. Nováček, P.Y. Vandenbussche, Regularizing knowledge graph embeddings via equivalence and inversion axioms, in: Machine Learning and Knowledge Discovery in Databases, Springer, 2017, pp. 668–683.
[56] S. Guo, Q. Wang, L. Wang, B. Wang, L. Guo, Jointly embedding knowledge graphs and logical rules, in: Conference on Empirical Methods in Natural Language Processing, 2016, pp. 192–202.
[57] B. Ding, Q. Wang, B. Wang, L. Guo, Improving knowledge graph embedding using simple constraints, 2018, arXiv preprint arXiv:1805.02408.
[58] Z. Cao, Q. Xu, Z. Yang, Q. Huang, ER: equivariance regularizer for knowledge graph completion, in: AAAI Conference on Artificial Intelligence, Vol. 36, No. 5, AAAI Press, 2022, pp. 5512–5520.
[59] S. Guo, Q. Wang, B. Wang, L. Wang, L. Guo, Semantically smooth knowledge graph embedding, in: Annual Meeting of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing, Vol. 1, 2015, pp. 84–94.
[60] N. Jain, T.K. Tran, M.H. Gad-Elrab, D. Stepanova, Improving knowledge graph embeddings with ontological reasoning, in: International Semantic Web Conference, Springer, 2021, pp. 410–426.
[61] N. Hubert, P. Monnin, A. Brun, D. Monticolo, Treat different negatives differently: Enriching loss functions with domain and range constraints for link prediction, in: Extended Semantic Web Conference, Springer, 2024, pp. 22–40.