Knowledge Graph Embeddings: Open Challenges and Opportunities

Russa Biswas[1]
Hasso-Plattner-Institut, Germany

Lucie-Aimée Kaffee
Hasso-Plattner-Institut, Germany

Michael Cochez
Vrije Universiteit Amsterdam, Netherlands
Elsevier Discovery Lab, Netherlands

Stefania Dumbrava
ENSIIE, France

Theis Jendal
Aalborg University, Denmark

Matteo Lissandrini
Aalborg University, Denmark

Vanessa Lopez
IBM Research Dublin, Ireland

Eneldo Loza Mencía
research.eneldo.net, Germany

Heiko Paulheim
Universität Mannheim, Germany

Harald Sack
FIZ Karlsruhe, Germany
Karlsruhe Institute of Technology, AIFB, Germany

Edlira Kalemi Vakaj
Birmingham City University, UK

Gerard de Melo
Hasso-Plattner-Institut, Germany
University of Potsdam, Germany

Abstract

While Knowledge Graphs (KGs) have long been used as valuable sources of structured knowledge, in recent years, KG embeddings have become a popular way of deriving numeric vector representations from them, for instance, to support knowledge graph completion and similarity search. This study surveys advances as well as open challenges and opportunities in this area. For instance, the most prominent embedding models focus primarily on structural information. However, there has been notable progress in incorporating further aspects, such as semantics, multi-modal, temporal, and multilingual features. Most embedding techniques are assessed using human-curated benchmark datasets for the task of link prediction, neglecting other important real-world KG applications. Many approaches assume a static knowledge graph and are unable to account for dynamic changes. Additionally, KG embeddings may encode data biases and lack interpretability. Overall, this study provides an overview of promising research avenues to learn improved KG embeddings that can address a more diverse range of use cases.

[1] corresponding author

Transactions on Graph Data and Knowledge, Vol. 1, Issue 1, Article No. 42, pp. 42:1–42:31
Transactions on Graph Data and Knowledge
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

2012 ACM Subject Classification Computing methodologies → Semantic networks

Keywords and phrases Knowledge Graphs, KG embeddings, Link prediction, KG applications

Digital Object Identifier 10.4230/TGDK.1.1.42

Category Survey

Funding *Michael Cochez*: Partially funded by the Graph-Massivizer project, funded by the Horizon Europe programme of the European Union (grant 101093202).

Received Date of submission **Accepted** Date of acceptance **Published** Date of publishing

Editor TGDK section area editor

1 Introduction

A Knowledge Graph (KG) is a semantic network that organises knowledge in a graph using entities, relations, and attributes. It captures semantic relationships and connections between entities, allowing for rapid searching, reasoning, and analysis. KGs are directed labelled graphs that can represent a variety of structured knowledge across a wide range of domains including e-commerce [96, 130], media [137], and life science [23], to name a few. They enable the integration of structured knowledge from diverse sources, laying the groundwork for applications such as question-answering systems, recommender systems, semantic search, and information retrieval. Google [155], eBay [130], Amazon [96], and Uber [58] are examples of companies that have developed in-house enterprise KGs for commercial purposes, which are not publicly available. The term "Knowledge Graph" was first used in the literature in 1972 [149] and later revived by Google in 2012 with the introduction of the Google KG. Broad-coverage open KGs, such as DBpedia [11], Freebase [19], YAGO [158], and Wikidata [173], are either developed using heuristics, manually curated, or automatically or semi-automatically extracted from structured data.

While the structured knowledge in KGs can readily be used in many applications, *KG embeddings* open up new possibilities. A KG embedding encodes semantic information and structural relationships by representing entities and relations in a KG as dense, low-dimensional numeric vectors. This entails developing a mapping between entities and relations and vector representations that accurately capture their characteristics and relationships.

KG embeddings allow for effective computation, reasoning, and analysis, while maintaining semantics and structural patterns. Link prediction and KG completion are perhaps the most well-known uses of KG embeddings. Although KGs store vast amounts of data, they are often incomplete. For instance, given the KG in Figure 1, which is an excerpt from DBpedia, it will not be possible to answer the questions:

Q1*: Where is Berkshire located?*, and

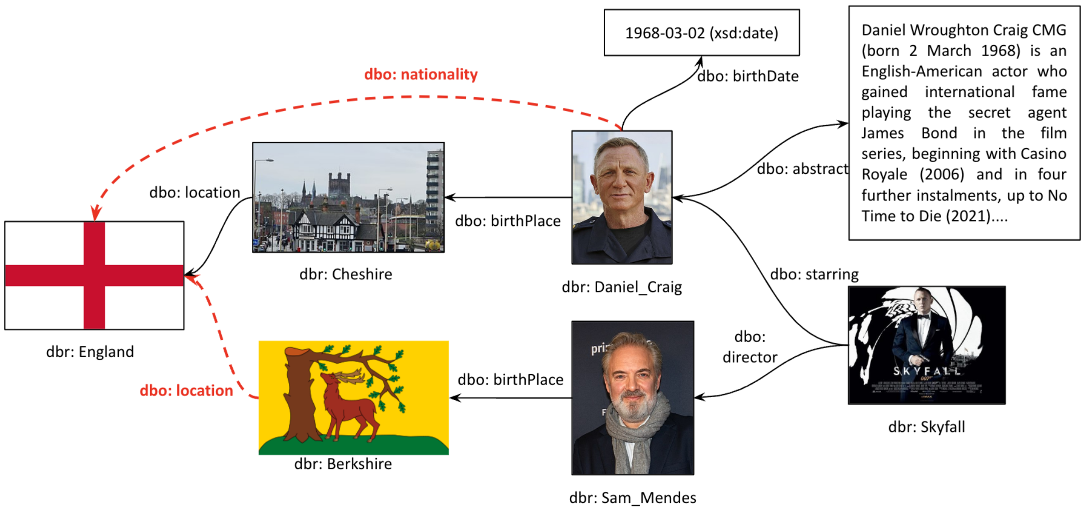
Q2*: What is the nationality of Daniel Craig?*

Responding to Q1 requires the prediction of the missing entity in the triple <dbr:Berkshire[2], dbo:locatedIn, ?>. Similarly, for Q2, one would need to infer the nationality of Daniel Craig from the information available in the KG. The effectiveness of KG-based question-answering applications may therefore be enhanced by using embeddings to predict the missing links in a KG. This is referred to as KG completion.

Other applications of KG embeddings include similarity search, entity classification, recommender systems, semantic search, and question answering. Additionally, an embedding converts

[2] For example, we will often shorten the IRIs using prefixes. For example, in dbr:Berkshire, dbr: stands for `http://dbpedia.org/resource/`, and hence the identifier is a shorthand for `http://dbpedia.org/resource/Berkshire`. Simlarly, dbo: stands for `http://dbpedia.org/ontology/`.

Figure 1 Excerpt from DBpedia, with red dashed lines representing possible inferred relations.



symbolic knowledge into numerical representations, making it possible to incorporate structured knowledge into machine learning and AI models, enabling reasoning across KGs.

Although prominent KG embedding models are widely used across diverse applications, there is potential to learn improved embeddings addressing an even broader range of input information and opening up new opportunities. For instance, one can account for additional signals in the KG beyond the structural information, such as multi-modal and hierarchical information, as well as external textual data, or information related to a certain domain or context. Some models struggle to adequately represent rare or long-tail entities, while others are unable to cope with little or no training data. Additionally, there is potential to design models that better account for dynamic and temporal information in the KG. Likewise, KGs are often multilingual, which may enable improved representations. Some models have trouble capturing asymmetric links as well as complex relationships such as hierarchical, compositional, or multi-hop relationships. The bias in KGs may also be reflected in the corresponding embeddings. Most models also lack explicit interpretability or explainability. This paper focuses on describing the relevant research addressing the aforementioned KG embedding models' inadequacies and then discussing the untapped areas for future research.

The rest of the paper is organised as: Section 2 gives an overview of the definitions and notations related to KGs, followed by Section 3 summarising mainstream KG embedding models. Next, Section 4 provides an overview of models that exploit additional kinds of information often neglected by traditional KG embedding models, along with a discussion of remaining open challenges. Section 5 sheds some light on important application areas of KG embeddings. Finally, Section 6 concludes the paper with a discussion and an outlook of future work.

2 Preliminaries

This section provides formal definitions and relevant notational conventions used in this paper.

▶ **Definition 1** (Knowledge Graph). *A KG $\mathcal{G}$ is a labelled directed graph, which can be viewed as a set of knowledge triples $\mathcal{T} \subseteq \mathcal{E} \times \mathcal{R} \times (\mathcal{E} \cup \mathcal{L})$, where $\mathcal{E}$ is the set of nodes, corresponding to entities (or resources), $\mathcal{R}$ is the set of relation types (or properties) of the entities, and $\mathcal{L}$ is the set of literals. An entity represents a real-world object or an abstract concept. Often the labels of entities and relations are chosen to be URIs or IRIs (Internationalised Resource Identifiers).*

▶ **Definition 2** (Triple). *Given a KG $\mathcal{G}$, we call $(e_h, r, e_t) \in \mathcal{T}$ a triple, where $e_h \in \mathcal{E}$ is the subject, $r \in \mathcal{R}$ is the relation, and $e_t \in \mathcal{E} \cup \mathcal{L}$ is the object. The subject is also called the head entity, and an object $e_t \in \mathcal{E}$ may be referred to as the tail entity. Triples with literals as objects, i.e., $e_t \in \mathcal{L}$ are known as attributive triples. In this paper, we use the notation $<e_h,r,e_t>$, with angle brackets, to indicate a triple.*

Relations (or Properties): Depending on the nature of the objects in a triple, one may distinguish two main kinds of relations:

- Object Relation (or Property), in which an entity is linked to another entity. For instance, in the triple <*dbr:Daniel_Craig*, *dbo:birthPlace*, *dbr:Cheshire*>, *dbr:Daniel_Craig* and *dbr:Cheshire* are head and tail entities, respectively, and *dbo:birthPlace* is an Object Relation (or Property).
- Data Type Relation (or Property), in which the entity is linked to a literal. For instance, we find the date *"1868-03-02"* in the triple <*dbr:Daniel_Craig*, *dbo:birthDate*, *"1868-03-02"*>, and therefore the relation *dbo:birthDate* is a Data Type Relation (or Property).

Additionally, an entity e can also be linked to classes or semantic types of the entity. For example, DBpedia uses *rdf:type* as r, while Freebase uses *isA*. A triple of the form $<e$, *rdf:type*, $\mathcal{C}_k>$ hence implies that $e \in \mathcal{E}$ is an entity, $\mathcal{C}_k \in \mathcal{C}$ is a class, $\mathcal{C}$ is the set of semantic types or classes, and e is an instance of $\mathcal{C}_k$. Often, the semantic types or the classes in a KG are organised in a hierarchical tree structure. An entity may belong to more than one class.

Literals: A KG can have many types of literal values and examples of common attribute types are as follows:

- *Text literals*: These store information in the form of free natural language text and are often used for labels, entity descriptions, comments, titles, and so on.
- *Numeric literals*: Dates, population sizes, and other data saved as integers, real numbers, etc. provide valuable information about an entity in a KG.
- *Image literals*: These literals can, for example, be used to store a visual representation of the entity, but can also contain the outcome of a medical scan, or a chart.

It is also possible that there is additional information (such as video or audio) stored external to the graph. The graph can then contain an IRI or other kind of identifier that references the external resource, its location, or both.

3 Knowledge Graph Embeddings

KG embedding models represent entities and relationships in a KG in a low-dimensional vector space for various downstream applications. A typical KG embedding model is characterised by the following aspects, as detailed by Ji et al. [82]: (1) The *Representation Space* may be a single standard Euclidean vector space, separate Euclidean vector spaces for entities and relations, or matrices, tensors, multivariate Gaussian distributions, or mixtures of Gaussians. Some methods also use complex vectors or hyperbolic space to better account for the properties of relationships. (2) A *scoring function* serves to represent relationships by quantifying the plausibility of triples in the KG, with higher scores for true triples and lower scores for false/negative/corrupted ones. (3) *Encoding models* are responsible for learning the representations by capturing relational

Table 1 Categorisation of classic Knowledge Graph Embedding Models

Categories	Models
Translational Models	TransE [20] and its variants, RotatE [160], etc,
Gaussian Embeddings	KG2E [66], TransG [192]
Semantic Matching Models	RESCAL [124] and its extensions, DistMult [198], HoIE [123], SME [21]
Neural Network Models	NTN [156] , HypER [14], ConvE [37], ConvKB [31]
Graph Neural Networks	GCN [92], R-GCN [148], GraphSAGE [60], GAT [172], KGAT [179], ComplEx-KG [170], SimlE [90]
Path-based Models	GAKE [43], PTransE [112], PTransR, RSN, PConvKB [83], RDF2vec [141]

interactions between entities. This is typically achieved by solving optimisation problems, often using factorisation approaches or neural networks. (4) *Auxiliary Information* in the KG may be incorporated, e.g., literals. This leads to enriched entity embeddings and relations, forming an ad-hoc scoring function integrated into the general scoring function.

An overview of different types of KG embedding models is given in Table 1. In the following, we explain each of these in more detail.

- **Translation**-based models use distance-based scoring functions to measure the plausibility of a fact as the distance between two entities. There are numerous variants. TransE [21] represents entities and relations as vectors in the same space, while TransH [184] introduces relation-specific hyperplanes. TransR [114] uses relation-specific spaces but requires a projection matrix for each relation. TransD [80] simplifies TransR by using two vectors for each entity-relation pair. TranSparse [81] employs two separate models, TranSparse(share) and TranSparse(separate), to modify projection vectors or matrices without considering other aspects. TransA [84] replaces the traditional Euclidean distance with the Mahalanobis distance, demonstrating better adaptability and flexibility as an indicator for performance improvement.
- **Gaussian Embeddings**: KG2E [66] and TransG [192] are probabilistic embedding models that incorporate uncertainty into their representation. KG2E uses multi-Gaussian distributions to embed entities and relations, representing the mean and covariance of each entity or relation in a semantic feature space. TransG, in contrast, uses a Gaussian mixture model to represent relations, addressing multiple relationship semantics and incorporating uncertainty. Both models offer unique approaches to representing entities and relations.
- **Semantic Matching** models rely on the notion of semantic similarity to define their scoring function. These include tensor decomposition models such as RESCAL, a tensor factorisation model that represents entities and relations as latent factors [124], capturing complex interactions between them. DistMult [199] simplifies the scoring function of RESCAL by using diagonal matrices, leading to more efficient computations. SimplIE [90] is a simpler model that uses a rule-based approach to extract relations from sentences. RotatE [161] introduces rotational transformations to model complex relationships in KGs. ComplEx [170] extends DistMult by introducing complex-valued embeddings, enabling it to capture both symmetric and antisymmetric relations. HolE [124] employs circular correlation to capture compositional patterns in KGs. TuckER [14] is a linear model based on Tucker decomposition of the binary tensor representation of triples.
- **Neural network** based models draw on the powerful representation learning abilities of modern deep learning. Neural Tensor Networks (NTN) [156] allow mediated interaction of entity vectors via a tensor. ConvE [37] uses 2D convolutions over embeddings to predict missing links in KGs. ConvKB [31] represents each triple as a 3-column matrix and applies

convolution filters to generate multiple feature maps, which are concatenated into a single feature vector. This vector is multiplied with a weight vector to produce a score, used for predicting the validity of the triple. HypER [14] generates convolutional filter weights for each relation using a hyper-network approach.

- **Graph Neural Network** models are neural networks that operate directly on the graph structure, often with information propagation along edges. GCN [92] and GraphSAGE [60] are graph convolutional techniques that combine information from neighbouring nodes in a graph to enable efficient learning of node representations in large-scale graphs. R-GCN [148] extends GCN to handle different relationships between entities in graph-structured data using a CNN model to learn hidden layer representations that encode local network structure and node attributes, growing linearly with the number of graph edges. GAT [172] employs an attention mechanism to dynamically allocate weights to neighbouring nodes, focusing on salient neighbours and capturing expressive representations. KGAT [179] applies the concept of graph attention networks to KG embeddings, taking into account entity and relation information, as well as capturing complicated semantic linkages and structural patterns. ComplEx-KG [170] is a complex-valued embedding-based extension of ComplEx, a bilinear model for KG embeddings. SimplE [90] uses a simplified scoring function for large KGs that is scalable and optimised for efficiency.
- **Path**-based models such as PTransE [112] and PTransR [113] represent entities and relations in the KG as vectors and learn embeddings based on relation-specific translation operations along edge paths. RSN [203] models the KG as a recursive structure, aggregating embeddings of connected entities and capturing structural information through recursive path-based reasoning. PConvKB [83] extends the ConvKB model and uses an attention mechanism on the paths to measure the local importance in relation paths. GAKE [43] is a graph-aware embedding model that takes into consideration three forms of graph structure: neighbour context, path context, and edge context. RDF2Vec [141] uses random walks over the graph structure to generate node and edge sequences, which are then used as input for training word2vec skip-gram models, which yield entity and relation embeddings.

Traditional KG embedding methods primarily take into account the triple information but neglect other potentially valuable signals encountered in KGs, such as multimodality, temporality, multilinguality, and many more. Additionally, these models often assume KGs are static in nature and have cold-start problems when incorporating new entities and relations. Also, real-world KGs often exhibit sparsity, noisiness, and bias, which may adversely affect embedding models.

4 Opportunities and Challenges

KG embeddings are widely used to capture semantic meaning and enable improved comprehension, reasoning, and decision-making across a diverse range of applications. However, the traditional KG embedding models described earlier neglect a series of important opportunities and aspects. In the following, in Section 4.1, we consider auxiliary information that may be present in KGs but is often neglected in KG embeddings, e.g., multimodal, multilingual, and dynamic knowledge. Subsequently, in Section 4.2, we discuss further more general issues, such as bias and explainability. Recent research has made notable progress in addressing these issues. The remainder of the section summarises pertinent recent research along with a discussion of open research challenges.

4.1 Auxiliary Information

Prominent KG embedding models such as those enumerated in Section 3 focus primarily on the structure of the KG, i.e., on structural information pertaining to entities and their relationships. To improve the latent representations of entities and relations, new lines of research attempt to draw on additional forms of information present in the KG. This section offers an overview of existing research in this regard, along with discussions of relevant shortcomings and recommendations for further research.

4.1.1 Multimodal KG Embeddings

Many approaches for representation learning on entities and relations ignore the variety of data modalities in KGs. In a Multimodal KG (MKG), entities and attributes of these entities may have different modalities, each providing additional information about the entity. An effective learned representation captures correspondences between modalities for accurate predictions, as described by Gesese et al. [53]. The used modalities depend on the application area, but can include text, images, numerical, and categorical values. Inductive approaches are required for modelling MKGs that encompass a variety of data modalities, as assuming that all entities have been observed during training is impractical. Learning a distinct vector for each entity and using enumeration for all possible attribute multimodal values to predict links is usually infeasible.

- **Text:** One of the early approaches for text extends TransE by incorporating word2vec SkipGram and training a probabilistic version in the same embedding space, anchoring via Freebase entities and the word embedding model vocabulary [183]. This enables link prediction for previously unknown entities. Relations are treated without differentiation of types. A combination of DistMult and CNN [169] tackles this issue by modelling the textual relations via dependency paths extracted from the text. Other models such as DKRL [194] and Jointly (BOW) [196] use the word2vec Continuous Bag-Of-Words (CBOW) approach to encode keywords extracted from textual entity descriptions, while Text Literals in KGloVe [30] uses these in combination with the graph context to train a GloVe model. However, the alignment between KG and word model is achieved using string matching and therefore struggles with ambiguous entity names. Veira et al. [171] use Wikipedia articles to construct relation-specific weighted word vectors (WWV). Convolutional models, such as DKRL (CNN) [194] and RTKRL [65], use word order to represent relations, considering implicit relationships between entities. Multi-source Knowledge Representation Learning (MKRL) [164] uses position embedding and attention in CNNs to find the most important textual relations among entity pairs. STKRL [188] extracts reference sentences for each entity and treats the entity representation as a multi-instance learning model. Recurrent neural models such as Entity Descriptions-Guided Embedding (EDGE) [178] and Jointly (ALSTM) [196] use attention-based LSTMs with a gating mechanism to encode entity descriptions, capturing long-term relational dependencies. The LLM encoder BERT is used in Pretrain-KGE [212] to generate initial entity embeddings from entity descriptions and relations, and subsequently feed them into KG embedding models for final embeddings. Other research uses LLMs [16, 181, 120, 3] to produce representations at word, sentence, and document levels, merging them with graph structure embeddings. KG-BERT [?] optimises the BERT model on KGs, followed by KG-GPT2 [17] fine-tuning the GPT-2 model. MTL-KGC [91] enhances the effectiveness of KG-BERT by combining prediction and relevance ranking tasks. Saxena et al. [147] similarly transform the link prediction task into a sequence-to-sequence problem by verbalizing triplets into questions and answers, overcoming the scalability issues of KG-BERT. Masked Language Modeling (MLM) has been introduced to encode KG text, with MEMKGC [28] predicting masked entities using the MEM classification model. StAR [174] uses bi-encoder-style textual encoders for text along with a scoring module, while SimKGC leverages bi-encoding for the textual encoder. LP-BERT [104] is a hybrid method that combines MLM Encoding for pre-training with LLM and Separated Encoding for fine-tuning.

- **Numeric literals** are addressed by several prominent models. MT-KGNN [166] trains a relational network for triple classification and an attribute network for attribute value regression, focusing on data properties with non-discrete literal values. KBLRN [50] combines relational, latent, and numerical features using a probabilistic PoE method. LiteralE [97] incorporates literals into existing latent feature models for link prediction, modifying the scoring function and using a learnable transformation function. TransEA [190] has two component models: a new attribute embedding model and a translation-based structure embedding model, TransE. These embedding approaches, however, fail to fully comprehend the semantics behind literal and unit data types. Additionally, most models lack proper mechanisms to handle multi-valued literals.
- **Image and Video** models account for multimedia content. There is a large body of work on visual relationship detection, i.e., identifying triples portrayed in visual content, using datasets such as VisualGenome [95] and methods such as VTransE [207]. IKLR [193] enriches KG embeddings by retrieving images for each entity from ImageNet. The respective set of pre-trained image embeddings is subsequently combined by an attention-based multi-instance learning method into a joint representation space of entities and relations. This additionally enables identifying the most relevant images for each entity.
- **General multi-modal** KG embedding models may be used both for better link prediction between existing entities and to impute missing values. One approach [128] combines different neural encoders to learn embeddings of entities and multimodal evidence types used to predict links. Then, DistMult or ConvE is employed to produce a score reflecting the probability that a triple is correct. In addition, neural decoders are applied over the learned embeddings to generate missing multimodal attributes, such as numerical values, text and images, from the information in the KG. Moreover, decoders can be invoked to generate entity names, descriptions, and images for previously unknown entities. A blueprint for multimodal learning from KGs is introduced by Ektefaie et al. [40]. Graph methods are employed to combine different datasets and modalities while leveraging cross-modal dependencies through geometric relationships. Graph Neural Networks (GNN) are used to capture interactions in multimodal graphs and learn a representation of the nodes, edges, subgraphs, or entity graph, based on message-passing strategies. Multimodal graphs find increasing application not only in computer vision and language modelling but also in natural sciences and biomedical networks [105], as well as in physics-informed GNNs that integrate multimodal data with mathematical models [154].

Limitations: Some of the key challenges reported in the literature that require further attention include: (1) Utilising multimodal information and multimodal fusion (from two or more modalities) to perform a prediction (e.g, classification, regression, or link prediction), even in the presence of missing modalities [128, 100, 40, 33]. (2) Modality collapse, that is when only a subset of the most helpful modalities dominates the training process. The model may overly rely on that subset of modalities and disregard information from the others that may be informative. This can be due to an imbalance in the learning process or insufficient data for one or more modalities and it can lead to sub-optimal representations [40]. (3) Generalisation across domains, modalities, and transfer learning of embeddings across different downstream tasks. In general, there is a high variance in the performance of multimodal methods [128, 109]. (4) Developing multimodal imputation models that are capable of generating missing multimodal values. While research in MKGs has predominantly focused on language (text) and vision (images) modalities, there is a need to explore multimodal research in other modalities and domains as well [128]. (5) Robustness to noise and controlling the flow of information within MKGs from more accurate predictions. While multimodal triples provide more information, not all parts of this additional data are necessarily informative for all prediction downstream tasks [100, 70, 128]. (6) Efficient and scalable frameworks that can handle the complexity during training and inference [33, 109]. Large KGs are challenging for all embedding-based link prediction techniques, and multimodal embeddings are not significantly worse because they can be viewed as having additional triples. However, multimodal encoder/decoders are more expensive to train [128] and techniques for batching and sampling are usually required for training. By addressing these challenges, we can unlock the full potential of MKGs and advance our understanding in various domains.

4.1.2 Schema/Ontology Insertion in KG Embeddings

While many real-world KGs come with schemas and ontologies, which may be rich and expressive, this does not hold for many of the benchmark datasets used in the evaluation of KG embeddings, in particular in the link prediction field. Therefore, the use of ontological knowledge for improving embeddings has drawn comparatively little attention.

In a very recent survey [208], the authors have reviewed approaches that combine ontological knowledge with KG embeddings. The authors distinguish between *pre* methods (methods applied before training the embedding), *joint* (during training of the embedding), and *post* (after training the embedding) methods. In their survey, joint methods are the most common approaches, usually incorporating the ontological knowledge in the loss function [10, 25, 39, 38, 51, 56, 98, 113, 143, 194, 205]. In such approaches, loss functions of existing KGE models are typically altered in a way such that ontologically non-compliant predictions are penalised. This is in line with a recent proposal of evaluation functions that not only take into account the ranking of correct triples but also the ontological compliance of predictions [74]. Some approaches also foresee the parallel training of class encoders [194] or class embeddings [64] to optimise the entity embeddings.

Pre methods observed in the literature come in two flavours. The first family of approaches exploit ontologies by inferring implicit knowledge in a preprocessing step and embedding the resulting graph enriched with inferred knowledge [75, 143]. The second family of approaches exploits ontologies in the process of sampling negative triples, implementing a sampling strategy that has a higher tendency to create ontologically compliant (and thus harder) negative examples [10, 57, 77, 98, 194], or builds upon adversarial training setups [116].

The *post* methods in the aforementioned survey are actually modifications of the downstream task, not the embedding method, and thus do not affect the embedding method per se.

The fact that most approaches fall into the *joint* category also limits them by being bound to one single embedding model, instead of being universally applicable. At the same time, most approaches have a very limited set of schema or ontology constraints they support (e.g., only domains and ranges of relations), while general approaches that are able to deal with the full spectrum of ontological definitions, or even more complex expressions such as SHACL constraints, remain very rare.

4.1.3 Relation Prediction Models

Relation prediction in KGs is a fundamental task that involves predicting missing or unobserved relations (properties) between entities in a KG. For instance, in Figure 1, relation prediction aims to predict the relation *dbo:starring* between entities *dbr:Daniel_Craig* and *dbr:Skyfall*.

Some of the classical KG embedding models such as translational models, and semantic matching models are often also used to predict missing relations. However, one of the pioneer models that focused on improving the relation prediction task is ProjE [153]. The model projected entity candidates onto a target vector representing input data, using a learnable combination operator to avoid transformation matrices followed by an optimised ranking loss of candidate entities. CNN-based models, in contrast, are argued to obtain richer and more expressive feature embeddings compared to traditional approaches. Attention-based embeddings enhance this approach further by capturing both entity and relation features in any given context or multihop neighbourhood [118]. Prior research on relation prediction, which was restricted to encyclopaedic KGs alone, disregarded the rich semantic information offered by lexical KGs, which resulted in the issue of shallow understanding and coarse-grained analysis for knowledge acquisition. HARP [182] extends earlier work by proposing a hierarchical attention module that integrates multiple semantic signals, combining structured semantics from encyclopaedic KGs and concept semantics from lexical KGs to improve relation prediction accuracy.

Self-supervised training objectives for multi-relational graph representation have as well given promising results. This may be achieved using a simplistic approach by incorporating relation prediction into the commonly used 1-vs-All objective [27]. The previously mentioned path-based embedding models may also be used, but often overlook sequential information or limited-length entity paths, leading to the potential loss of crucial information. GGAE [106] is a novel global graph attention embedding network model that incorporates long-distance information from multi-hop paths and sequential path information for relation prediction. The effectiveness of KG embedding models for relation prediction is typically assessed using rank-based metrics, which evaluate the ability of models to give high scores to ground-truth entities.

Limitations: Although embedding-based models for relation prediction in KGs have advanced significantly, they have several shortcomings. (1) Most of the models struggle to capture transitivity, which is essential for understanding relations that change over time or apply in different contexts. (2) They also struggle to handle rare relations, which can result in biased predictions. (3) Although embedding techniques are intended to accommodate multi-relational data, capturing complex interactions between numerous relations remains challenging. (4) KGs can contain relations with different semantic heterogeneity. For example, imagine a KG with a relation called *hasPartner* that represents any type of close partnership, such as business partners or friends. This relationship is semantically different from *hasSpouse*. Relation prediction models are often unable to distinguish between such relations with related but different meanings. (5) Relation prediction models provide limited support for temporal and contextual information. Temporal information, however, is handled by the temporal KG embedding models presented in Section 4.1.5.

4.1.4 Hierarchical and N-to-M Modeling in KG Embeddings

Crucial to the success of using KG embeddings for link prediction is their ability to model relation connectivity patterns, such as symmetry, inversion, and composition. However, many existing models make deterministic predictions for a given entity and relation and hence struggle to adequately model N-to-M relationships, where a given entity can stand in the same relationship to many other entities, as for instance for the *hasFriend* relationship [121].

A particular important case is that of hierarchical patterns, which, albeit ubiquitous, still pose significant challenges. Indeed, modelling them with knowledge embeddings often requires additional information regarding the hierarchical typing structure of the data [194] or custom techniques [211, 210], as discussed next.

Various approaches have been proposed for modelling hierarchical structures. Li et al. [107] proposes a joint embedding of entities and categories into a semantic space, by integrating structured knowledge and taxonomy hierarchies from large-scale knowledge bases, as well as a Hierarchical Category Embedding (HCE) model for hierarchical classification. This model additionally incorporates the ancestor categories of the target entity when predicting context entities, to capture the semantics of hierarchical concept category structures.

Another method used for hierarchical modelling centres around the usage of clustering algorithms [211]. The authors define a three-layer hierarchical relation structure (HRS) for KG relation clusters, relations, and subrelations. Based on this, they extend classic translational embedding models to learn better knowledge representations. Their model defines the embedding of a knowledge triple based on the sum of the embedding vectors for each of the HRS layers.

The Type-embodied Knowledge Representation Learning (TKRL) [194] model uses entity-type information in KG embeddings to model hierarchical relations. Following the TransE approach, relations are translated between head and tail KG entities in the embedding space. For each entity type, type-specific projection matrices are built using custom hierarchical type encoders, projecting the heads and tails of entities into their type spaces.

Limitations: Although they intend to better represent the structure of a KG, the limitations of such KG embeddings include: (1) It is challenging to model interactions that transcend numerous hierarchy levels, resulting in a limited ability to capture cross-hierarchy linkages. For instance, *Arnold Schwarzenegger* is an *actor*, a *film director* as well as a *politician*, leading to the entity belonging to different branches of the class hierarchy in the KG. (2) The depth of the hierarchy or branching factor of an n-to-m relationship can affect how effective the embeddings are, e.g., in very fine-grained or coarse-grained hierarchies, performance may suffer. (3) Training and inference with hierarchical embeddings can be computationally intensive, particularly in ultrafine-grained hierarchies.

4.1.5 Temporal KG Embeddings

Most KG completion methods assume KGs to be static, which can lead to inaccurate prediction results due to the constant change of facts over time. For instance, neglecting the fact that <*Barack Obama*, *presidentOf*, *USA*> only holds from 2009 to 2017 can become crucial for KG completion. Emerging approaches for Temporal Knowledge Graph Completion (TKGC) incorporate timestamps into facts to improve the result prediction. These methods consider the dynamic evolution of KGs by adding timestamps to convert triples into quadruples using several strategies [22]:

- **Tensor Decomposition** based models in KG completion transform a KG into a 3-dimensional binary tensor, with three modes representing head, relation, and tail entities to learn their corresponding representations by tensor decomposition. The addition of timestamps as an additional mode of tensor (4-way tensor) for TKGC allows for low-dimensional representations of timestamps for scoring functions. For TKGC, Canonical Polyadic (CP) decomposition is used on quadruple facts [111]. The authors employ an imaginary timestamp for static facts, while complex-valued representation vectors may be used for asymmetric relations [99]. Temporal smoothness penalties are used to ensure that neighbouring timestamps obtain similar representations. Multivector representations [195] are learned using CP decomposition, allowing the model to adjust to both point timestamps and intervals. A temporal smoothness penalty for timestamps is created and expanded to a more generic autoregressive model. Tucker decomposition can be used for TKGC [151], treating KGs as 4-way tensors and scoring functions that consider interactions among entities, relations, and timestamps, relaxing the requirement for identical embedding dimensions of entities, relations, and timestamps.
- **Timestamp-based Transformation** models involve generating synthetic time-dependent relations by concatenating relations with timestamps (e.g., *presidentOf:2009-2017*), converting <*Barack Obama*, *presidentOf*, *USA*> to <*Barack Obama*, *presidentOf:2009-2017*, *USA*> [101]. This however may lead to more synthetic relations than necessary. An improvement is to derive optimal timestamps for concatenating relations by splitting or merging existing time intervals [135]. The concatenation of relation and timestamp as a sequence of tokens is also provided as an input making the synthetic relation adaptive to different formats like points, intervals, or modifiers [49]. Others [177] argue that different relations rely on different time resolutions, such as a life span in years or a birth date in days. Multi-head self-attention is

adopted on the timestamp-relation sequence to achieve adaptive time resolution. In the TKGC model, timestamps are often considered linear transformations that map entities/relations to corresponding representations. The timestamps are also treated as hyperplanes, dividing time into discrete time zones [32]. An additional relational matrix is included to map entities to be relation-specific to improve expressiveness for multi-relational facts [185]. To capture dynamics between hyperplanes, a GRU may be applied to the sequence of hyperplanes [163]. Another approach [102] encodes timestamps into a one-hot vector representing various time resolutions, such as centuries or days to achieve time precision.

- **KG Snapshots** can be considered as a series of snapshots/subgraphs taken from a KG, with each subgraph holding facts labelled with a timestamp. Therefore, a temporal subgraph evolves with changing relation connections. The link prediction problem can be solved by utilising Markov models [197] to infer the multi-relational interactions among entities and relations over time and can be trained using a recursive model. Probabilistic entity representations based on variational Bayesian inference can be adopted to model entity features and uncertainty jointly [110]. The dynamic evolution of facts can be modelled using an autoregressive approach [85], incorporating local multi-hop neighbouring information and a multi-relational graph aggregator. Alternatively, a multilayer GCN can capture dependencies between concurrent facts with gated components to learn long-term temporal patterns [108]. Continuous-time embeddings can encode temporal and structural data from historical KG snapshots [63].
- **Historical Context** based models focus on the chronological order of facts in a KG, determined by the availability of timestamps, which enable predicting missing links by reasoning with the historical context of the query. An attention-based reasoning process has been proposed [62] as the expansion of a query-dependent inference subgraph, which iteratively expands by sampling neighbouring historical facts. Another approach uses path-based multi-hop reasoning by propagating attention using a two-stage GNN through the edges of the KG, using the inferred attention distribution [86]. The model captures displacements at two different granularities, i.e., past, present, and future and the magnitude of the displacement. Two heuristic-based tendency scores Goodness and Closeness [12] have been introduced to organise historical facts for link prediction. Historical facts are aggregated based on these scores, followed by a GRU for dynamic reasoning. It is observed that history often repeats itself in KGs [213], leading to the proposal of two modes of inference: Copy and Generation.

Limitations: Although recently many TKGC models have been proposed that resolve the issues of classical KG embedding models with timestamps, some intriguing possibilities for future studies on TKGC include: (1) External knowledge such as relational domain knowledge, entity types, and semantics of entities and relationships can be added to the limited structural/temporal information during model learning to enhance prediction accuracy. (2) Due to the time dimension and intricate relationships between facts and timestamps, time-aware negative sampling should be investigated in TKGC. (3) Most methods assume timestamps are available, while in some cases only relative time information is known. For example, we would know that a person lived in a city after they were born, but neither when the person was born, nor when they started living there. (4) With the constant evolution of the real-world KGs, TKGC should be regarded as an incremental or continual learning problem.

4.1.6 Dynamic KG Embeddings

As discussed in the previous section, incorporating timestamps is one way to handle changes; however, facts may be added, altered, or deleted over time, are not foreseen [94], and would typically require a complete re-computation of the embedding model. Such an approach might still be feasible for KGs like DBpedia, which have release cycles of weeks or months [69], but not for continuously updated KGs such as Wikidata, let alone examples of even more highly dynamic KGs, e.g., digital twins, which may continuously change every second. Moreover, naïvely recomputing embeddings for an only slightly changed KG may lead to drastic shifts in the embeddings of existing entities, e.g., due to stochastic training behaviour. This would require a recalibration of downstream models consuming those embeddings, as they would not be *stable* [187, 93].

While a few approaches for embedding dynamic graphs (not necessarily KGs) have been proposed [89], many of them focus on embedding a series of snapshots of KGs, rather than developing mechanisms for embedding a dynamic KG. Thus, they do not support *online learning*, i.e., continuously adjusting the KG embedding model whenever changes occur.

Approaches capable of online learning are much scarcer. One of the first was puTransE [165], which continuously learns new embedding spaces. Similarly, Wewer et al. [187] investigate updating the link prediction model by incorporating change-specific epochs forcing the model to update the embeddings related to added or removed entities and/or relations.

Embeddings based on random walks can be adapted to changes in the graph by extracting new walks around the changed areas [115], or by applying local changes to the corpus of random walks [146]. The latter approach also supports the deletion of nodes and edges. DKGE [189] learns embeddings using gated graph neural networks and requires retraining only vectors of affected entities in the online learning part. Similarly, OUKE first learns static embeddings and computes dynamic representations only locally using graph neural networks. The two representations are then combined into a dynamic embedding vector. The idea of only updating embeddings of affected entities is also pursued by RotatH [186]. A different strategy is considered by Navi [93], which learns a surrogate model to reconstruct the entity embeddings based on those of neighbouring existing entities. This surrogate model is then used to recompute the embedding vectors for new entities or entities with changed contexts.

Limitations: The main limitations in the existing approaches so far are threefold: (1) In most models, only addition to KGs is studied, while deletion is not the focus, an exception is the work by Wewer et al. [187].[3] (2) The stability of the resulting embeddings, which is crucial for downstream applications, has rarely been analysed systematically. (3) The applicability in a true real-time scenario, as it would be required, e.g., for digital twins, is unclear for most approaches, which are evaluated on snapshots.

4.1.7 Inductive KG Embedding

In the inductive setting, graph representation learning involves training and inference of partially or completely disjoint sets of nodes, edges, and possibly even relationships types. In practice, from the specific set of known structures, it tries to generalise knowledge that enables reasoning with unseen graph objects by exploiting information on the structures involving them and the data attached to them [46]. The case of link prediction involves being able to predict the existence of a link between two previously unseen nodes (head and tail) by reasoning about their connections to other known nodes (i.e., nodes observed during training) or by reasoning about their attributes (e.g., features similar to those of nodes seen during training).

Therefore, in the most common setting, relationship types do not change, but training involves a given KG and inference involves a completely or partially different graph. Overall, the crucial point is that there must be some form of shared information that allows for *inferring* a description

[3] Even for papers using different versions of public KGs e.g., DBpedia or YAGO, the majority of changes are additions, and most benchmarks used in the evaluation of the papers mentioned above, usually have much more additions than deletions.

of an unknown entity or edge from a small set of known attributes. For example, a common approach allows for predictions involving previously unseen, or out-of-sample, entities that attach to a known KG with a few edges adopting known relationship types [47]. In this case, a few nodes in the KG seen during training are used as anchors and called NodePieces. A full NodePiece vocabulary is then constructed from anchor nodes and relation types. Given a new node, an embedding representation is obtained using elements of the constructed NodePiece vocabulary extracting a *hash code* for it given by the sequence of k closest anchors, combined with discrete anchor distances, and a relational context connecting relations. Other approaches extract a local subgraph of one or more nodes and consider the structures within such a subgraph trying to learn an inductive bias able to infer entity-independent relational semantics [167]. This approach is then also adopted to predict missing facts in KGs, i.e., to predict a missing relation between two entities. Similarly, NBFNet [214] instead encodes the representation of a pair of nodes using the generalised sum of all path representations between the two nodes and with each path representation as the generalised product of the edge representations in the path. In this case, the operation is modelled along the line of a generalised Bellman-Ford algorithm that computes the shortest paths from a single source vertex to all of the other vertices by taking into account edge weights. Here, operators to compute the length of the shortest path are learned for a specific downstream task.

The aforementioned methods are designed for the case where the only information available are triples connecting entities and do not take into account node or edge properties. Conversely, when properties are taken into account, e.g., textual data describing entities, this information can be exploited as node or edge features. A typical case is that of networks that adopt an auto-encoder architecture to encode node representations and decode edges as a function over the representation of node pairs. Among those, GraphSAGE [60] was the first inductive GNN able to efficiently generate embeddings for unseen nodes by leveraging node features, e.g., textual attributes. Later methods, including BLP [35] create embeddings for entities by encoding the description with a language model fine-tuned on a link prediction objective. This model can then be used inductively, as long as nodes have a description.

Limitations: All these approaches have only scratched the surface of the need for KG embeddings. In particular, challenges persist in terms of (1) scalability, e.g., the possibility of learning inductive biases from small representative samples of the graph; (2) exploiting well-known feature extraction from graphs and KGs, as existing methods tend to disregard the possibility of using structural features, e.g., betweenness, page rank, relational neighbourhood and characteristic sets [122]; (3) moreover, while GNNs seem the most promising and expressive architecture, their ability to produce inductive relation aware KG representations are limited in their treatment of rich vocabularies of relation types (typically limited to fewer than a hundred), their ability to exploit information at more than 3 hops of distance, and the possibility to generate a representation for very sparse feature sets. Finally, known challenges that apply to transductive methods, e.g., distribution shift and how to update the model or decide to train it from scratch, still apply. Finally, the ability to work in an inductive fashion might increase the risk of data leakages, which already exist in non-inductive settings [41]. The use of GNNs that learn how to aggregate information from node and edge attributes raises more concerns when the training data involves private data; how to ensure that private data is not leaked through the model, e.g., via differentially private KG embedding [61], is still an open question.

4.1.8 Multilingual KG Embeddings

Providing multilingual information in a KG is crucial to ensure wide adoption across different language communities [87]. Languages in KGs can have different representations; e.g., in Wikidata, each entity has a language-independent identifier, and labels in different languages are indicated with the `rdfs:label` property [88]. Therefore, in Wikidata, entities do not need alignment across languages. In DBpedia, there is one entity per language, derived from the respective language Wikipedia [103]. Therefore, different language entities on the same concept can have different facts stated about them. Here, an alignment using the `owl:sameAs` property is necessary to ensure the different entities are connected across languages and enable seamless access to information for all language communities. The different representations of languages in the different KGs can heavily influence which way the KG can be embedded. For example, if provided with a KG per language as in DBpedia, different language KGs might be embedded separately and then aligned or can be fused for usage in downstream applications [73].

One of the downstream tasks of multilingual KG embeddings is KG completion. Finding new facts given machine-readable data such as a KG is a tedious task for human annotators, even more so when the graph covers a wide range of languages. Addressing these challenges, recent work has employed KG embeddings across languages to predict new facts in a KG.

One of the large challenges of multilingual KG embeddings is the knowledge inconsistency across languages, i.e., the vastly different number of facts per language. Fusing different languages to overcome such knowledge inconsistencies for multilingual KG completion can improve performance across languages, especially for lower-resourced languages [73]. To fuse different languages, KGs need to be aligned across languages. Such alignment can be done jointly with the task of multilingual KG completion [24, 168, 26].

Another approach for multilingual KG completion is leveraging large language models' (LLM) knowledge about the world to add new facts to a KG. As LLMs are not trained towards KG completion and are biased towards English, Song et al. [157] introduce global and local knowledge constraints to constrain the reasoning of answer entities and to enhance the representation of query context. Hence, the LLMs are better adapted for the task of multilingual KG completion.

Limitations: Although most of the existing multilingual KG embedding models focus on having a unified embedding space across different language versions of the KGs, these embeddings have several shortcomings. (1) The potential of the model to learn and generalise relations between entities in different languages is often restricted by sparse cross-lingual links, resulting in less accurate cross-lingual representations of entities. (2) Polysemy, which occurs when a word has numerous meanings, can be difficult to address across languages, resulting in ambiguity in cross-lingual representations. (3) Entities and relations can have very context-dependent and language-specific meanings, which is a challenging task for multilingual embeddings to capture the nuances of the context. (4) Resource imbalances may result in low-resource languages having inadequate training data and linguistic resources, impacting the entity and relation embeddings.

4.2 General Challenges

In addition to the goal of accounting for a broader spectrum of available information, there are more general challenges and opportunities for KG embedding models: (1) KG embedding models can inherit biases from training data, thereby reinforcing societal preconceptions. (2) Scalable embedding approaches are required for large-scale KGs with millions or billions of elements and relations. (3) Improving the interpretability and explainability of embeddings remains a challenge.

4.2.1 Bias in KG Embeddings

KGs, which serve as the foundation for KG embeddings, are regarded as crucial tools for organizing and presenting information, enabling us to comprehend the vast quantities of available data. Once constructed, KGs are commonly regarded as "gold standard" data sources that uphold the accuracy of other systems, thus making the objectivity and neutrality of the information they convey vital concerns. Biases inherent to KGs may become magnified and spread through KG-based systems [150]. Traditionally, bias can be defined as "*a disproportionate weight in favour of or against an idea or thing, usually in a way that is closed-minded, prejudicial, or unfair*"[4]. Taking into account the bias networking effect for KGs, it is crucial that various types of bias are already acknowledged and addressed during KG construction [78].

Biases within KGs, as well as the approaches to address them, differ from those found in linguistic models or image classification. KGs are sparse by nature, i.e., only a small number of triples are available per entity. In contrast, linguistic models acquire the meaning of a term through its contextual usage in extensive corpora, while image classification leverages millions of labelled images to learn classes. Biases in KGs can arise from various sources, including the design of the KG itself, the (semi-)automated generation of the source data, and the algorithms employed to sample, aggregate, and process the data. These source biases typically manifest in expressions, utterances, and textual sources, which can then permeate downstream representations and in particular KG embeddings. Additionally, we must also account for a wide range of human biases, such as reporting bias, selection bias, confirmation bias, overgeneralisation, and more.

Biases in KGs as the source of KG embeddings can arise from multiple sources. Data bias occurs already in the data collection process or simply from the available source data. Schema bias depends on the chosen ontology for the KG or simply is already embedded within the used ontologies [78]. Inferential bias might result from drawing inferences on the represented knowledge. Ontologies are typically defined by a group of knowledge engineers in collaboration with domain experts and consequently (implicitly) reflect the world views and biases of the development team. Ontologies are also prone to encoding bias depending on the chosen representation language and modeling framework. Moreover, biases in KG embeddings may in particular arise from the chosen embedding method as for instance induced by application-specific loss functions. Inferential biases, which may arise at the inferencing level, such as reasoning, querying, or rule learning, are mostly limited to KGs themselves and rarely propagate to KG embeddings. A simple example of inferencing bias might be the different SPARQL entailment regimes, which in consequence, might be responsible for different results that different SPARQL endpoints deliver despite containing the same KG [2, 54].

Collaboratively built KGs, such as DBpedia or GeoNames, also exhibit social bias, often arising from the western-centric world view of their main contributors [36]. In addition, some "truths" represented in such KGs may be considered controversial or opinionated, which underlines the importance of provenance information.

For KG embeddings that represent a vector space-based approximation of the structural and semantic information contained in a KG, one of the main sources of bias lies in the sparsity and incompleteness of most KGs. KG embeddings trained on incomplete KGs might favour entities for which more information is available [136]. Moreover, if the underlying KG is biased, then KG embeddings trained on this base data will as well be, and in fact bias may even be amplified. De-biasing of KG embeddings requires methods for detecting as well as removing bias in KG embeddings. Depending on the underlying embedding model, this task might become complex and requires finetuning of embeddings with respect to certain sensitive relations [44, 45, 9].

4.2.2 Reliability and Scalability of KG Embeddings

KG embedding methods suffer from many issues in terms of scalability. For example, many studies experiment mainly on (poorly constructed) subsets of Freebase and Wordnet, the infamous FB15k

---

[4] Wikipedia article on bias. `https://en.wikipedia.org/wiki/Bias`, retrieved 2023-11-28.

and WN18 [1], which are known to suffer from information leakage. These datasets contain in the order of a few million triples and rarely go beyond 1,000 relationship types, usually focusing on subgraphs with 200 or fewer. Recently, more realistic datasets have been proposed in terms of the quality of the data involved and of the link prediction task adopted [145]. Nonetheless, even these are far from being representative of typical real-world KG applications. Consider that DBpedia contains 52M distinct triples involving 28M distinct literals and as many distinct entities, with 1.3K distinct relationship types. Indeed, a recent Wikidata snapshot contains 1.926 billion triples, involving more than 600M entities and 904M distinct literals across 9K relationship types [134]. The size of real-world KGs is far beyond the capabilities of current methods, and the current results on small controlled benchmarks cannot be seen as representative of their scalability and reliability on real-world deployment. This perhaps also suggests the need for methods designed end-to-end to consider cases where different models can be learned for different subgraphs and then combined in a modular fashion. Last but not least, as KG embedding methods are adopted for tasks that go beyond link prediction, e.g., KG alignment [159], we refer to the well-known issues of scale in terms of dataset size (number of triples) and in terms of heterogeneity (scale of the vocabulary of relationships and attributes), as well as to new important issues based on the number of KGs to align, i.e., scale in terms of the number of distinct KG sources [15].

4.2.3 Explainability of KG Embeddings

One of the persistent difficulties is the development of KG embedding methods to enhance interpretability and explainability. This includes comprehending the reasoning and decision-making processes of KG embedding models as well as providing explanations for their predictions. KG embeddings have several advantages over conventional representations produced by deep learning algorithms, including their absence of ambiguity and the ability to justify and explain decisions [125]. Additionally, they can offer a semantic layer to help applications such as question-answering, which are normally handled by text-based brute force techniques. CRIAGE [129] is one such tool that can be used to understand the impact of adding and removing facts. GNNExplainer [202] is proposed for the explainability of the predictions done by GNNs. Deep Knowledge-Aware Networks [176] and Knowledge-aware Path Recurrent Networks [180] have witnessed a surge in attention to recommendation systems. They model sequential dependencies that link users and items. OpenDialKG [117] is a corpus that aligns KGs with dialogues and presents an attention-based model that learns pathways from dialogue contexts and predicts relevant novel entities. These models offer a semantic and explicable layer for conversational agents and recommendations, aiding in the completion and interpretation of the predictions.

Limitations: However, there are still a number of limitations: (1) The lack of standardised evaluation standards makes it difficult to compare different approaches and assess performance consistently. (2) Improving interpretability often comes at the expense of performance and striking a balance between interpretability and performance still remains a challenge. (3) User-centric evaluation is necessary to understand the practical utility of explainable KG embeddings. (4) Current research on KG embedding explainability often focuses on global or model-level explanations, ignoring the importance of contextual and domain-specific explanations.

4.2.4 Complex Logical Query Answering and Approximate Answering of Graph Queries

The link prediction task is often seen as a graph completion task. However, it can equivalently be cast as a query-answering task for a very simple query. For example, if we predict the tail of the triple $<h,r,?>$, the task is equivalent to answering the corresponding query as if the graph had all the missing information. Recently, researchers started investigating how we could answer such queries if they are more complex, a task known as complex logical query answering[5]. The goal is, given a graph with missing information and a graph query, to produce the answers to the query as if the graph were complete (or more commonly, produce a ranking of possible answers).

One might naïvely assume that this can be solved by first completing the graph and then performing a traditional graph query on the completed graph. The issue is, however, that a very large KG can never be complete. This is because link prediction models do not yield a set of missing edges, but rather a ranking of possible completions for an incomplete triple.

We can distinguish three main lines of work in this area. The reader is referred to relevant surveys [138, 29] for more details. The first group of approaches are those that make use of a link predictor, like the ones introduced above. These methods *decompose* the query into triples and then use the link prediction model to make predictions for the triples. The first approach of this type was CQD [7], which uses fuzzy logic to combine the outputs of the link predictor. Further developments for this type of model include QTO [13], which materialises all intermediate scores for the link predictors and makes sure that edges existing in the graph are always regarded as more certain than those predicted by the link predictor. Another newer approach is Adaptive CQD [8], which improves CQD by calibrating the scores of the link predictor across different relation types.

A second group of approaches are referred to as projection approaches, and the earliest approaches in this domain are of this type. These methods are characterised by the restriction that they can only answer DAG-shaped graph queries. They are inspired by translation-based link predictors. Starting from the entities in the query (in this context called the anchors), they project them with a relation-specific model to a representation for the tail entity. This representation then replaces the other occurrences as a subject of the variable in the query. If a variable occurs in more than one object position, a model is invoked to combine the computed projections into a single representation (called the intersection). The first approach of this type was Graph Query Embedding (GQE) [59], which did the above using vectors as representations, simple linear projections, and an MLP with element-wise mean for the intersection. Later examples include Query2Box [139], which uses axis-aligned hyperplanes to represent the outcomes of projections and intersections, and BetaE [140], which instead uses the beta distribution.

A final group of approaches is message-passing-based. These are very flexible and can deal with more query shapes than the above. This method regards the query as a small graph and embeds that complete query into a single embedding. Then, answers to the query are found simply by retrieving the entities of which the embedding is close to that query in the embedded space. A notable example is MPQE [34], which uses a relational graph convolutional network (R-GCN) to embed the query. The flexibility of these models is illustrated by StarQE [4], which can even answer hyper-relational queries (very similar to RDF-star).

Limitations: As indicated in the survey by Ren et al. [138], there are still very many open questions in this domain. (1) One aspect is that current approaches only support small subsets of all possible graph queries. For example, hardly any work attempts to answer cyclic queries, queries with variables on the relation position, or only variables in the whole query. (2) Also, the graph formalism currently used is limited; only very few approaches can deal with literal data, and there is no word yet on temporal KGs or the use of background semantics.

[5] also sometimes approximate query answering, multi-hop reasoning, or query embedding

5 Applications

Recent research on KG embeddings has shown broad potential across diverse application domains such as search engines [42], recommendation systems [48], question-answering systems [72], biomedical and healthcare informatics [5], e-commerce [209], social network analysis [152], education [200], and scientific research [119]. However, in this study, we highlight two such domains: recommendation and biomedical/therapeutic use cases.

5.1 KG Embedding for Recommendation

Recommender systems (RSs) are an integral part of many online services and applications to provide relevant content and products tailored to their users. Many RSs identify user preference patterns assuming that users with similar past behaviour have similar preferences, e.g., people that watch the same movies are likely to do so also in the future, an approach commonly referred to as collaborative filtering [68, 67]. Yet, many existing methods only work in a warm-start setting, where it is assumed that all users and items have been seen during training [60, 204]. Moreover, methods that try to deal with cold-start settings, where for some users or items only user–item interactions are known and only at inference time [201, 204], making them unable to handle situations where this type of data is sparse, e.g., long-tail users and items. Therefore, we can see this problem as a link prediction problem, and we can also distinguish between a transductive setting and an inductive setting. In the transductive setting, some approaches try to exploit other contextual information from KGs, e.g., semantic annotations, taxonomies, item descriptions, or categories, to overcome these problems. In particular, a large body of methods exploits both domain-specific and open-domain KGs integrated with user and item information. In practice, users and items are nodes connected by special domain-specific relation types, e.g., a rating or a purchase, and item nodes are represented with additional connections to other entities describing their categories, features, producers, and provenance. This information, in the form of a Collaborative KG, is adopted as additional side information in the recommendation process [179, 175, 126]. These methods can be grouped into three categories:

1. path-based methods, which capture information from distant nodes but tend to dismiss much of the structural information in KG and are very dependent on the paths selected during training [**?**, 191, 162];
2. embedding-based methods, which use existing transductive graph embedding approaches to capture the semantic relations of the graph structure, such as TransR [205] or Node2Vec [55], further applying them in recommendation scenarios [126, 206]; and
3. structural-based methods, which use GNNs to aggregate structural information of each node's neighborhood [175, 179].

Among these, GNNs have recently shown promising results thanks to their ability to model relations and capture high-order connectivity information by combining KGs and collaborative data (user–item interactions) [179]. Nonetheless, these approaches often rely on transductive methods, making them unable to handle frequent changes in the graph. Moreover, their user–item representation often is limited to a single relation type and still cannot fully exploit the contextual knowledge offered by open-domain KGs, due to only very few relation types being considered. Furthermore, these approaches need to be able to exploit both the structure of the graph and the attributes describing the items.

5.2 Multimodal KG Embeddings for Biomedical and Therapeutic Use

In the biomedical domain, KGs are a natural way to model and represent complex biomedical structured data, such as molecular interactions, signalling pathways and disease co-morbidities [105]. Information from a single source usually does not provide sufficient data, and various state-of-the-art studies have shown that incorporating multiple heterogeneous knowledge sources and modalities yields better predictions [100, 52, 70]. Learning an effective representation that leverages the topology of these multimodal and heterogeneous KGs to create optimised embedding representations is key to applying AI models. These optimised embeddings can then be fed into link prediction models, such as for interactions between proteins [79], drugs [52], drug-targets [52, 100], or drug indication/contraindications for diseases [70].

For instance, Otter-Knowledge [100] uses MKGs built from diverse sources, where each node has a modality assigned, such as textual (e.g., protein function), numerical (e.g., molecule mass), categorical entities (e.g., protein family), and modalities for representing protein and molecules. For each modality in the graph, a model is assigned to compute initial embeddings, e.g., pre-trained language models such as ESM [142] and MolFormer [144] are used for protein sequences and molecules' SMILES, respectively. A GNN is then invoked to enrich the initial representations and train a model to produce knowledge-enhanced representations for drug molecules and protein entities. These representations can improve drug-target binding affinity prediction tasks [71], even in the presence of entities not encountered during training or having missing modalities.

During training, attribute modalities are treated as relational triples of structured knowledge instead of predetermined features, making them first-class citizens of the MKG [128, 100]. The advantage of this approach is that entity nodes are not required to carry all multimodal properties or project large property vectors with missing values. Instead, the projection is done per modality and only when such a modality exists for the entity.

6 Discussion and Conclusion

Currently, the vast majority of evaluations of knowledge graph embeddings are conducted on the task of link prediction. At the same time, embeddings created with such techniques are used across a wide range of diverse downstream tasks, such as recommender systems, text annotation and retrieval, fact validation, data interpretation and integration, to name just a few. This raises the question: *How suitable is the effectiveness of a link prediction task as a predictor of the applicability of a particular KGE method for a particular downstream task?*

While the evaluation of link prediction is quite standardised with respect to benchmark datasets and evaluation metrics, the field of downstream applications is much more diverse and less standardised. Some frameworks, such as GEval [127] and kgbench [18], offer a greater variety of tasks and evaluations, including evaluation metrics and dataset splits.

Some studies have looked into characterizing the representation capabilities of different KGE methods. They, for instance, analyse whether different classes are separated in the embedding space [6, 76, 215]. More recently, the DLLC benchmark [132] has been proposed, which allows for analysing which types of classification problems embeddings produced by a particular method can address. Other studies analyse the distance function in the resulting embedding spaces, finding that while most approaches create embedding spaces that encode entity similarity, others focus on entity relatedness [131], and that some methods can actually be altered to focus more on similarity and relatedness [133].

In addition, link prediction, entity categorisation, KG completion, and KG embeddings are crucial for a number of downstream activities, such as entity recommendation, relation extraction, question-answering, recommender systems, semantic search, and information retrieval. Models that leverage user profiles, historical interactions, and KGs can deliver personalised recommendations, capture similarity and relevance, and increase accuracy and relevance. KG embeddings also improve the accuracy of relation extraction by adding structured knowledge. The majority of existing KG embedding models are generalised, that is, they are trained and evaluated on open KGs for KG completion. However, task-specific KG embeddings would be quite advantageous in various kinds of applications, which still remains an open research task. They can be optimised for creating representations for specific tasks, improving performance, focusing on relevant information extraction, resolving data scarcity, and thereby improving interpretability and explainability. With the use of domain-specific data or constraints, these embeddings can be trained to grasp and reason about the relationships and semantics unique to that domain.

Recent ongoing research also reveals that when KG embeddings and LLMs are combined, a symbiotic relationship results, maximising the benefits of each methodology. While LLMs help to integrate textual knowledge, improve entity and relation linking, promote cross-modal fusion, and increase the explainability of KG embeddings, KG embeddings provide structured knowledge representations that improve the contextual comprehension and reasoning of LLMs. Therefore, future research may focus on building more robust and comprehensive models for knowledge representation, reasoning, and language understanding as a result of these interrelated effects.

KG embeddings will continue to evolve and serve an important role in enabling effective knowledge representation, reasoning, and decision-making as KGs grow in scale and complexity. Advances in KG embeddings offer the ability to make it easier to convert unstructured data into structured knowledge, reveal deeper insights, and enhance intelligent applications, as highlighted in this study.

References

1 Farahnaz Akrami, Mohammed Samiul Saeef, Qingheng Zhang, Wei Hu, and Chengkai Li. Realistic re-evaluation of knowledge graph completion methods: An experimental study. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, pages 1995–2010. ACM, 2020. doi:10.1145/3318464.3380599.

2 Mehwish Alam, George. Fletcher, Antonie. Isaac, Aidan. Hogan, Diana. Maynard, Heiko. Paulheim, Harald. Sack, Elena. Simperl, Lise. Stork, and Hideaki. van Erp, Marieke. amd Takeda. Bias in knowledge graph systems. In *Proc. of Dagstuhl Seminar 22372: Knowledge Graphs and their Role in the Knowledge Engineering of the 21st Century, 2022; Vol. 12(9)*, pages 106–109. Dagstuhl Reports, 2022.

3 Mirza Mohtashim Alam, Md Rashad Al Hasan Rony, Mojtaba Nayyeri, Karishma Mohiuddin, MST Mahfuja Akter, Sahar Vahdati, and Jens Lehmann. Language model guided knowledge graph embeddings. *IEEE Access*, 10:76008–76020, 2022. doi:10.1109/access.2022.3191666.

4 Dimitrios Alivanistos, Max Berrendorf, Michael Cochez, and Mikhail Galkin. Query embedding on hyper-relational knowledge graphs. In *10th International Conference on Learning Representations, ICLR 2020, virtual, April 25-29, 2022*. OpenReview.net, 2022. URL: https://openreview.net/forum?id=4rLw09TgRw9.

5 Mona Alshahrani, Maha A Thafar, and Magbubah Essack. Application and evaluation of knowledge graph embeddings in biomedical data. *PeerJ Computer Science*, 7:e341, 2 2021. doi:10.7717/peerj-cs.341.

6 Faisal Alshargi, Saeedeh Shekarpour, Tommaso Soru, and Amit Sheth. Concept2vec: Metrics for evaluating quality of embeddings for ontological concepts. *arXiv preprint arXiv:1803.04488*, abs/1803.04488, 2018. doi:10.48550/arXiv.1803.04488.

7 Erik Arakelyan, Daniel Daza, Pasquale Minervini, and Michael Cochez. Complex query answering with neural link predictors. In *International Conference on Learning Representations*. OpenReview.net, 2021. URL: https://openreview.net/forum?id=Mos9F9kDwkz.

8 Erik Arakelyan, Pasquale Minervini, Daniel Daza, Michael Cochez, and Isabelle Augenstein. Adapting neural link predictors for data-efficient complex query answering. 1 2023. arXiv:2301.12313.

9 Mario Arduini, Lorenzo Noci, Federico Pirovano, Ce Zhang, Yash Raj Shrestha, and Bibek Paudel. Adversarial learning for debiasing knowledge graph embeddings. *ArXiv*, abs/2006.16309, 6 2020. doi:10.48550/arXiv.2006.16309.

10 Siddhant Arora, Srikanta Bedathur, Maya Ramanath, and Deepak Sharma. IterefinE: Iterative KG refinement embeddings using symbolic knowledge. In *Automated Knowledge Base Construction*, 2020. URL: https://openreview.net/forum?id=fCQvGMT57w, doi:10.24432/C5NP46.

11 Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. DBpedia: A nucleus for a web of open

data. In *The semantic web*, volume 4825, pages 722–735. Springer Berlin Heidelberg, 2007. doi:10.1007/978-3-540-76298-0_52.

12 Luyi Bai, Xiangnan Ma, Mingcheng Zhang, and Wenting Yu. TPmod: A tendency-guided prediction model for temporal knowledge graph completion. *ACM Transactions on Knowledge Discovery from Data*, 15(3):1–17, 6 2021. doi:10.1145/3443687.

13 Yushi Bai, Xin Lv, Juanzi Li, and Lei Hou. Answering complex logical queries on knowledge graphs via query computation tree optimization. *arXiv preprint arXiv:2212.09567*, abs/2212.09567, 12 2022. URL: http://arxiv.org/pdf/2212.09567, arXiv:2212.09567, doi:10.48550/arxiv.2212.09567.

14 Ivana Balažević, Carl Allen, and Timothy M Hospedales. Hypernetwork knowledge graph embeddings. In *Proceedings of the International Conference on Artificial Neural Networks*, volume 11731, pages 553–565. Springer International Publishing, 2019. doi:10.1007/978-3-030-30493-5_52.

15 Matthias Baumgartner, Daniele Dell'Aglio, Heiko Paulheim, and Abraham Bernstein. Towards the web of embeddings: Integrating multiple knowledge graph embedding spaces with FedCoder. *Journal of Web Semantics*, 75:100741, 2023. doi:https://doi.org/10.1016/j.websem.2022.100741.

16 Russa Biswas, Harald Sack, and Mehwish Alam. MADLINK: Attentive multihop and entity descriptions for link prediction in knowledge graphs. *Semantic Web*, pages 1–24, 2022. doi:10.3233/sw-222960.

17 Russa Biswas, Radina Sofronova, Mehwish Alam, and Harald Sack. Contextual language models for knowledge graph completion. In *MLSMKG*, volume 2997. CEUR-WS.org, 2021. URL: https://ceur-ws.org/Vol-2997/paper3.pdf.

18 Peter Bloem, Xander Wilcke, Lucas van Berkel, and Victor de Boer. kgbench: A collection of knowledge graph datasets for evaluating relational and multimodal machine learning. In *European Semantic Web Conference*, volume 12731, pages 614–630. Springer International Publishing, 2021. doi:10.1007/978-3-030-77385-4_37.

19 Kurt D. Bollacker, Robert P. Cook, and Patrick Tufts. Freebase: A shared database of structured general human knowledge. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence*, pages 1962–1963. AAAI Press, 7 2007.

20 Antoine Bordes, Nicolas Usunier, Alberto García-Durán, Jason Weston, and Oksana Yakhnenko. Translating embeddings for modeling multi-relational data. *Advances in neural information processing systems*, 26:2787–2795, 2013. URL: https://proceedings.neurips.cc/paper/2013/hash/1cecc7a77928ca8133fa24680a88d2f9-Abstract.html.

21 Antoine Bordes, Jason Weston, Ronan Collobert, and Yoshua Bengio. Learning structured embeddings of knowledge bases. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, volume 25, pages 301–306. Association for the Advancement of Artificial Intelligence (AAAI), 2011. doi:10.1609/aaai.v25i1.7917.

22 Borui Cai, Yong Xiang, Longxiang Gao, He Zhang, Yunfeng Li, and Jianxin Li. Temporal knowledge graph completion: A survey. *arXiv preprint arXiv:2201.08236*, abs/2201.08236:6545–6553, 1 2022. doi:10.24963/ijcai.2023/734.

23 Alison Callahan, Jose Cruz-Toledo, Peter Ansell, and Michel Dumontier. Bio2RDF release 2: improved coverage, interoperability and provenance of life science linked data. In *Extended semantic web conference*, volume 7882, pages 200–212. Springer, Springer Berlin Heidelberg, 2013. doi:10.1007/978-3-642-38288-8_14.

24 Soumen Chakrabarti, Harkanwar Singh, Shubham Lohiya, Prachi Jain, and Mausam. Joint completion and alignment of multilingual knowledge graphs. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 11922–11938. Association for Computational Linguistics, 2022. doi:10.18653/v1/2022.emnlp-main.817.

25 Kai-Wei Chang, Wen-tau Yih, Bishan Yang, and Christopher Meek. Typed tensor decomposition of knowledge bases for relation extraction. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1568–1579. Association for Computational Linguistics, 2014. doi:10.3115/v1/d14-1165.

26 Xuelu Chen, Muhao Chen, Changjun Fan, Ankith Uppunda, Yizhou Sun, and Carlo Zaniolo. Multilingual knowledge graph completion via ensemble knowledge transfer. In *Findings of the Association for Computational Linguistics: EMNLP 2020, Online Event, 16-20 November 2020*, volume EMNLP 2020 of *Findings of ACL*, pages 3227–3238. Association for Computational Linguistics, 2020. doi:10.18653/v1/2020.findings-emnlp.290.

27 Yihong Chen, Pasquale Minervini, Sebastian Riedel, and Pontus Stenetorp. Relation prediction as an auxiliary training objective for improving multi-relational graph representations. In *3rd Conference on Automated Knowledge Base Construction, AKBC 2021, Virtual, October 4-8, 2021*, 2021. doi:10.24432/C54K5W.

28 Bonggeun Choi, Daesik Jang, and Youngjoong Ko. MEM-KGC: Masked entity model for knowledge graph completion with pre-trained language model. *IEEE Access*, 9:132025–132032, 2021. doi:10.1109/access.2021.3113329.

29 Michael Cochez, Dimitrios Alivanistos, Erik Arakelyan, Max Berrendorf, Daniel Daza, Mikhail Galkin, Pasquale Minervini, Mathias Niepert, and Hongyu Ren. Approximate answering of graph queries, 8 2023. to apear in the Compendium of Neuro-Symbolic Artificial Intelligence. URL: https://arxiv.org/pdf/2308.06585, doi:10.48550/arxiv.2308.06585.

30 Michael Cochez, Martina Garofalo, Jérôme Lenßen, and Maria Angela Pellegrino. A first experiment on including text literals in KGloVe. In *Joint proceedings of the 4th Workshop on Semantic Deep Learning (SemDeep-4) and NLIWoD4: Natural Language Interfaces for the Web of Data (NLIWOD-4) and 9th Question Answering over Linked Data challenge (QALD-9) co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, California, United States of America, October 8th*

- *9th, 2018*, volume 2241 of *CEUR Workshop Proceedings*, pages 103–106. CEUR-WS.org, 2018.
31 Tu Dinh Nguyen Dai Quoc Nguyen, Dat Quoc Nguyen, and Dinh Phung. A novel embedding model for knowledge base completion based on convolutional neural network. In *Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 327–333. Association for Computational Linguistics, 2018. doi:10.18653/v1/n18-2053.
32 Shib Sankar Dasgupta, Swayambhu Nath Ray, and Partha Talukdar. HyTE: Hyperplane-based temporally aware knowledge graph embedding. In *Proceedings of the 2018 conference on empirical methods in natural language processing*, pages 2001–2011. Association for Computational Linguistics, 2018. doi:10.18653/v1/d18-1225.
33 Daniel Daza, Dimitrios Alivanistos, Payal Mitra, Thom Pijnenburg, Michael Cochez, and Paul Groth. BioBLP: A modular framework for learning on multimodal biomedical knowledge graphs. *arXiv preprint arXiv:2306.03606*, 2023. doi:https://doi.org/10.48550/arXiv.2306.03606.
34 Daniel Daza and Michael Cochez. Message passing query embedding. In *ICML Workshop - Graph Representation Learning and Beyond*, 2020. doi:10.48550/arXiv.2002.02406.
35 Daniel Daza, Michael Cochez, and Paul Groth. Inductive entity representations from text via link prediction. In *Proceedings of the Web Conference 2021*, pages 798–808. ACM, 2021. doi:10.1145/3442381.3450141.
36 Gianluca Demartini. Implicit bias in crowdsourced knowledge graphs. In *Companion Proceedings of The 2019 World Wide Web Conference*, page 624–630. Association for Computing Machinery, 2019. doi:10.1145/3308560.3317307.
37 Tim Dettmers, Pasquale Minervini, Pontus Stenetorp, and Sebastian Riedel. Convolutional 2D knowledge graph embeddings. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, pages 1811–1818. Association for the Advancement of Artificial Intelligence (AAAI), 2018. doi:10.1609/aaai.v32i1.11573.
38 Boyang Ding, Quan Wang, Bin Wang, and Li Guo. Improving knowledge graph embedding using simple constraints. *arXiv preprint arXiv:1805.02408*, pages 110–121, 2018. doi:10.18653/v1/p18-1011.
39 Claudia d'Amato, Nicola Flavio Quatraro, and Nicola Fanizzi. Injecting background knowledge into embedding models for predictive tasks on knowledge graphs. In *European Semantic Web Conference*, volume 12731, pages 441–457. Springer, Springer International Publishing, 2021. doi:10.1007/978-3-030-77385-4_26.
40 Yasha Ektefaie, George Dasoulas, Ayusa Noori, Maha Farhat, and Marinka Zitnik. Multimodal learning with graphs. *Nature Machine Intelligence*, 5:340—350, 2023. doi:10.1038/s42256-023-00624-6.
41 Michael Ellers, Michael Cochez, Tobias Schumacher, Markus Strohmaier, and Florian Lemmerich. Privacy attacks on network embeddings. *CoRR*, abs/1912.10979, 2019. doi:https://doi.org/10.48550/arXiv.1912.10979.
42 Shahla Farzana, Qunzhi Zhou, and Petar Ristoski. Knowledge graph-enhanced neural query rewriting. In *Companion Proceedings of the ACM Web Conference 2023*, pages 911–919. ACM, 4 2023. doi:10.1145/3543873.3587678.
43 Jun Feng, Minlie Huang, Yang Yang, and Xiaoyan Zhu. GAKE: Graph aware knowledge embedding. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 641–651. ACL, 2016.
44 Joseph Fisher. Measuring social bias in knowledge graph embeddings. *ArXiv*, abs/1912.02761, 12 2019. doi:10.48550/arXiv.1912.02761.
45 Joseph Fisher, Arpit Mittal, Dave Palfrey, and Christos Christodoulopoulos. Debiasing knowledge graph embeddings. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7332–7345. Association for Computational Linguistics, 2020. doi:10.18653/v1/2020.emnlp-main.595.
46 Mikhail Galkin, Max Berrendorf, and Charles Tapley Hoyt. An open challenge for inductive link prediction on knowledge graphs. *arXiv preprint arXiv:2203.01520*, abs/2203.01520, 3 2022. URL: http://arxiv.org/pdf/2203.01520, doi:10.48550/arxiv.2203.01520.
47 Mikhail Galkin, Etienne Denis, Jiapeng Wu, and William L. Hamilton. NodePiece: Compositional and parameter-efficient representations of large knowledge graphs. In *International Conference on Learning Representations*. OpenReview.net, 2022. URL: https://openreview.net/forum?id=xMJWUKJnFSw.
48 Min Gao, Jian-Yu Li, Chun-Hua Chen, Yun Li, Jun Zhang, and Zhi-Hui Zhan. Enhanced multitask learning and knowledge graph-based recommender system. *IEEE Transactions on Knowledge and Data Engineering*, 35:10281–10294, 2023. doi:10.1109/tkde.2023.3251897.
49 Alberto García-Durán, Sebastijan Dumancic, and Mathias Niepert. Learning sequence encoders for temporal knowledge graph completion. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4816–4821. Association for Computational Linguistics, Association for Computational Linguistics, 2018. doi:10.18653/v1/d18-1516.
50 Alberto García-Durán and Mathias Niepert. KBLRN: End-to-end learning of knowledge base representations with latent, relational, and numerical features. pages 372–381, 2018.
51 Dinesh Garg, Shajith Ikbal, Santosh K Srivastava, Harit Vishwakarma, Hima Karanam, and L Venkata Subramaniam. Quantum embedding of knowledge for reasoning. *Advances in Neural Information Processing Systems*, 32:5595–5605, 2019. URL: https://proceedings.neurips.cc/paper/2019/hash/cb12d7f933e7d102c52231bf62b8a678-Abstract.html.
52 Aryo Pradipta Gema, Dominik Grabarczyk, Wolf De Wulf, Piyush Borole, Javier Antonio Alfaro, Pasquale Minervini, Antonio Vergari, and Ajitha Rajan. Knowledge graph embeddings in the biomedical domain: Are they useful? a look at link

prediction, rule learning, and downstream polypharmacy tasks. *arXiv.org*, abs/2305.19979, 5 2023. URL: `https://arxiv.org/pdf/2305.19979`, `arXiv:2305.19979`, `doi:10.48550/arxiv.2305.19979`.

53 Genet Asefa Gesese, Russa Biswas, Mehwish Alam, and Harald Sack. A survey on knowledge graph embeddings with literals: Which model links better literal-ly? *Semantic Web*, 12:617–647, 2021. `doi:10.3233/sw-200404`.

54 Paul Groth, Elena Simperl, Marieke van Erp, and Denny Vrandečić. Knowledge graphs and their role in the knowledge engineering of the 21st century (Dagstuhl seminar 22372). *Dagstuhl Reports*, 12:60–120, 2023. `doi:10.4230/DagRep.12.9.60`.

55 Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864. ACM, 2016. `doi:10.1145/2939672.2939754`.

56 Shu Guo, Quan Wang, Lihong Wang, Bin Wang, and Li Guo. Jointly embedding knowledge graphs and logical rules. In *Proceedings of the 2016 conference on empirical methods in natural language processing*, pages 192–202. Association for Computational Linguistics, 2016. `doi:10.18653/v1/d16-1019`.

57 Shu Guo, Quan Wang, Lihong Wang, Bin Wang, and Li Guo. Knowledge graph embedding with iterative guidance from soft rules. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, pages 4816–4823. Association for the Advancement of Artificial Intelligence (AAAI), 2018. `doi:10.1609/aaai.v32i1.11918`.

58 Ferras Hamad, Issac Liu, and Xian Xing Zhang. Food discovery with uber eats: Building a query understanding engine. uber engineering blog. 2018. Accessed: 2022-07-03.

59 William L. Hamilton, Payal Bajaj, Marinka Zitnik, Dan Jurafsky, and Jure Leskovec. Embedding logical queries on knowledge graphs. In *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pages 2030–2041, 2018. URL: `https://proceedings.neurips.cc/paper/2018/hash/ef50c335cca9f340bde656363ebd02fd-Abstract.html`.

60 William L Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 1025–1035, 2017. URL: `https://proceedings.neurips.cc/paper/2017/hash/5dd9db5e033da9c6fb5ba83c7a7ebea9-Abstract.html`.

61 Xiaolin Han, Daniele Dell'Aglio, Tobias Grubenmann, Reynold Cheng, and Abraham Bernstein. A framework for differentially-private knowledge graph embeddings. *Journal of Web Semantics*, 72:100696, 2022. `doi:https://doi.org/10.1016/j.websem.2021.100696`.

62 Zhen Han, Peng Chen, Yunpu Ma, and Volker Tresp. Explainable subgraph reasoning for forecasting on temporal knowledge graphs. In *International Conference on Learning Representations*, 2020.

63 Zhen Han, Zifeng Ding, Yunpu Ma, Yujia Gu, and Volker Tresp. Learning neural ordinary equations for forecasting future links on temporal knowledge graphs. In *Proceedings of the 2021 conference on empirical methods in natural language processing*, pages 8352–8364. Association for Computational Linguistics, 2021. `doi:10.18653/v1/2021.emnlp-main.658`.

64 Junheng Hao, Muhao Chen, Wenchao Yu, Yizhou Sun, and Wei Wang. Universal representation learning of knowledge bases by jointly embedding instances and ontological concepts. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1709–1719. ACM, 2019. `doi:10.1145/3292500.3330838`.

65 Ming He, Xiangkun Du, and Bo Wang. Representation learning of knowledge graphs via fine-grained relation description combinations. *IEEE Access*, 7:26466–26473, 2019. `doi:10.1109/access.2019.2901544`.

66 Shizhu He, Kang Liu, Guoliang Ji, and Jun Zhao. Learning to represent knowledge graphs with gaussian embedding. In *Proceedings of the 24th ACM international on conference on information and knowledge management*, pages 623–632. ACM, 2015. `doi:10.1145/2806416.2806502`.

67 Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. LightGCN: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 639–648. ACM, 2020. `doi:10.1145/3397271.3401063`.

68 Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*, pages 173–182. International World Wide Web Conferences Steering Committee, 2017. `doi:10.1145/3038912.3052569`.

69 Nicolas Heist, Sven Hertling, Daniel Ringler, and Heiko Paulheim. Knowledge graphs on the web-an overview. *Knowledge Graphs for eXplainable Artificial Intelligence*, pages 3–22, 2020. `doi:10.3233/ssw200009`.

70 Kexin Huang, Payal Chandak, Qianwen Wang, Shreyas Havaldar, Akhil Vaid, Jure Leskovec, Girish Nadkarni, Benjamin Glicksberg, Nils Gehlenborg, and Marinka Zitnik. Zero-shot prediction of therapeutic use with geometric deep learning and clinician centered design. *medRxiv*, 2023. `doi:10.1101/2023.03.19.23287458`.

71 Kexin Huang, Tianfan Fu, Wenhao Gao, Yue Zhao, Yusuf Roohani, Jure Leskovec, Connor W. Coley, Cao Xiao, Jimeng Sun, and Marinka Zitnik. Artificial intelligence foundation for therapeutic science. *Nature chemical biology*, 18:1033–1036, 2022. `doi:10.1038/s41589-022-01131-2`.

72 Xiao Huang, Jingyuan Zhang, Dingcheng Li, and Ping Li. Knowledge graph embedding based question answering. In *Proceedings of the twelfth ACM international conference on web search and data mining*, pages 105–113. ACM, 2019. `doi:10.1145/3289600.3290956`.

73 Zijie Huang, Zheng Li, Haoming Jiang, Tianyu Cao, Hanqing Lu, Bing Yin, Karthik Subbian, Yizhou Sun, and Wei Wang. Multilingual knowledge graph completion with self-supervised adaptive graph alignment. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, volume abs/2203.14987, pages 474–485. Association for Computational Linguistics, 2022. doi:10.18653/v1/2022.acl-long.36.

74 Nicolas Hubert, Pierre Monnin, Armelle Brun, and Davy Monticolo. Knowledge graph embeddings for link prediction: Beware of semantics! In *DL4KG@ ISWC 2022: Workshop on Deep Learning for Knowledge Graphs, held as part of ISWC 2022: the 21st International Semantic Web Conference*, volume 3342. CEUR-WS.org, 2022.

75 Andreea Iana and Heiko Paulheim. More is not always better: The negative impact of A-box materialization on RDF2vec knowledge graph embeddings. In *Workshop on Combining Symbolic and Subsymbolic Methods and their Applications (CSSA)*, volume abs/2009.00318. CEUR-WS.org, 9 2020.

76 Nitisha Jain, Jan-Christoph Kalo, Wolf-Tilo Balke, and Ralf Krestel. Do embeddings actually capture knowledge graph semantics? In *The Semantic Web: 18th International Conference, ESWC 2021, Virtual Event, June 6–10, 2021, Proceedings 18*, volume 12731, pages 143–159. Springer, Springer International Publishing, 2021. doi:10.1007/978-3-030-77385-4_9.

77 Nitisha Jain, Trung-Kien Tran, Mohamed H Gad-Elrab, and Daria Stepanova. Improving knowledge graph embeddings with ontological reasoning. In *International Semantic Web Conference*, volume 12922, pages 410–426. Springer, Springer International Publishing, 2021. doi:10.1007/978-3-030-88361-4_24.

78 Krzysztof Janowicz, Bo Yan 0003, Blake Regalia, Rui Zhu, and Gengchen Mai. Debiasing knowledge graphs: Why female presidents are not like female popes. In *Proceedings of the ISWC 2018 Posters & Demonstrations, Industry and Blue Sky Ideas Tracks co-located with 17th International Semantic Web Conference (ISWC 2018), Monterey, USA, October 8th - to - 12th, 2018*, volume 2180 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2018. URL: http://ceur-ws.org/Vol-2180/ISWC_2018_Outrageous_Ideas_paper_17.pdf.

79 Kanchan Jha, Sriparna Saha, and Snehanshu Saha. Prediction of protein-protein interactions using deep multi-modal representations. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 7 2021. doi:10.1109/IJCNN52387.2021.9533478.

80 Guoliang Ji, Shizhu He, Liheng Xu, Kang Liu, and Jun Zhao. Knowledge graph embedding via dynamic mapping matrix. In *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing*, pages 687–696. Association for Computational Linguistics, 2015. doi:10.3115/v1/p15-1067.

81 Guoliang Ji, Kang Liu, Shizhu He, and Jun Zhao. Knowledge graph completion with adaptive sparse transfer matrix. In *Proceedings of the Thirtieth AAAI conference on artificial intelligence*, volume 30, pages 985–991. Association for the Advancement of Artificial Intelligence (AAAI), 2016. doi:10.1609/aaai.v30i1.10089.

82 Shaoxiong Ji, Shirui Pan, Erik Cambria, Pekka Marttinen, and S Yu Philip. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(2):494–514, 2021.

83 Ningning Jia, Xiang Cheng, and Sen Su. Improving knowledge graph embedding using locally and globally attentive relation paths. In *European Conference on Information Retrieval*, volume 12035, pages 17–32. Springer International Publishing, 2020. doi:10.1007/978-3-030-45439-5_2.

84 Yantao Jia, Yuanzhuo Wang, Hailun Lin, Xiaolong Jin, and Xueqi Cheng. Locally adaptive translation for knowledge graph embedding. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, volume 30, pages 992–998. Association for the Advancement of Artificial Intelligence (AAAI), 2016. doi:10.1609/aaai.v30i1.10091.

85 Woojeong Jin, Meng Qu, Xisen Jin, and Xiang Ren. Recurrent event network: Autoregressive structure inferenceover temporal knowledge graphs. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6669–6683. Association for Computational Linguistics, 2020. doi:10.18653/v1/2020.emnlp-main.541.

86 Jaehun Jung, Jinhong Jung, and U Kang. Learning to walk across time for interpretable temporal knowledge graph completion. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 786–795. ACM, 2021. doi:10.1145/3447548.3467292.

87 Lucie-Aimée Kaffee. *Multilinguality in knowledge graphs*. PhD thesis, University of Southampton, 2021.

88 Lucie-Aimée Kaffee, Alessandro Piscopo, Pavlos Vougiouklis, Elena Simperl, Leslie Carr, and Lydia Pintscher. A glimpse into babel: An analysis of multilinguality in wikidata. In *Proceedings of the 13th International Symposium on Open Collaboration, OpenSym 2017, Galway, Ireland, August 23-25, 2017*, pages 14:1–14:5. ACM, 2017. doi:10.1145/3125433.3125465.

89 Seyed Mehran Kazemi, Rishab Goel, Kshitij Jain, Ivan Kobyzev, Akshay Sethi, Peter Forsyth, and Pascal Poupart. Representation learning for dynamic graphs: A survey. *The Journal of Machine Learning Research*, 21:2648–2720, 2020. URL: http://jmlr.org/papers/v21/19-447.html.

90 Seyed Mehran Kazemi and David Poole. Simple embedding for link prediction in knowledge graphs. *Advances in neural information processing systems*, 31:4289–4300, 2018.

91 Bosung Kim, Taesuk Hong, Youngjoong Ko, and Jungyun Seo. Multi-task learning for knowledge graph completion with pre-trained language models. In *COLING*, pages 1737–1743. International Committee on Computational Linguistics, 2020. doi:10.18653/v1/2020.coling-main.153.

92 Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks.

arXiv preprint arXiv:1609.02907, abs/1609.02907, 2016. doi:10.48550/arXiv.1609.02907.

93 Franz Krause. Dynamic knowledge graph embeddings via local embedding reconstructions. In *The Semantic Web: ESWC 2022 Satellite Events: Hersonissos, Crete, Greece, May 29–June 2, 2022, Proceedings*, volume 13384, pages 215–223. Springer, 2022. doi:10.1007/978-3-031-11609-4_36.

94 Franz Krause, Tobias Weller, and Heiko Paulheim. On a generalized framework for time-aware knowledge graphs. In *Towards a Knowledge-Aware AI: SEMANTiCS 2022—Proceedings of the 18th International Conference on Semantic Systems, 13-15 September 2022, Vienna, Austria*, volume 55, page 69. IOS Press, IOS Press, 2022. doi:10.3233/ssw220010.

95 Ranjay Krishna, Yuke Zhu, Oliver Groth, Justin Johnson, Kenji Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li-Jia Li, David A Shamma, et al. Visual genome: Connecting language and vision using crowdsourced dense image annotations. *International Journal of Computer Vision*, 123:32–73, 2017. doi:10.1007/s11263-016-0981-7.

96 Arun Krishnan. Making search easier: How amazon's product graph is helping customers find products more easily. amazon blog. 2018. Accessed: 2022-07-03.

97 Agustinus Kristiadi, Mohammad Asif Khan, Denis Lukovnikov, Jens Lehmann, and Asja Fischer. Incorporating literals into knowledge graph embeddings. In *Proceedings of the International Semantic Web Conference*, volume 11778, pages 347–363. Springer International Publishing, 2019. doi:10.1007/978-3-030-30793-6_20.

98 Denis Krompaß, Stephan Baier, and Volker Tresp. Type-constrained representation learning in knowledge graphs. In *International semantic web conference*, volume 9366, pages 640–655. Springer, Springer International Publishing, 2015. doi:10.1007/978-3-319-25007-6_37.

99 Timothée Lacroix, Guillaume Obozinski, and Nicolas Usunier. Tensor decompositions for temporal knowledge base completion. In *International Conference on Learning Representations*, 2019.

100 Hoang Thanh Lam, Marco Luca Sbodio, Marcos Martinez Gallindo, Mykhaylo Zayats, Raul Fernandez-Diaz, Victor Valls, Gabriele Picco, Cesar Berrospi Ramis, and Vanessa Lopez. Otter-knowledge: benchmarks of multimodal knowledge graph representation learning from different sources for drug discovery. *arXiv.org*, abs/2306.12802, 6 2023. URL: http://arxiv.org/pdf/2306.12802, arXiv:todo, doi:10.48550/arxiv.2306.12802.

101 Julien Leblay and Melisachew Wudage Chekol. Deriving validity time in knowledge graph. In *Companion Proceedings of the The Web Conference 2018*, pages 1771–1776. ACM Press, 2018. doi:10.1145/3184558.3191639.

102 Julien Leblay, Melisachew Wudage Chekol, and Xin Liu. Towards temporal knowledge graph embeddings with arbitrary time precision. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 685–694. ACM, 2020. doi:10.1145/3340531.3412028.

103 Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick Van Kleef, Sören Auer, et al. Dbpedia–a large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic web*, 6(2):167–195, 2015. doi:10.3233/sw-140134.

104 Da Li, Ming Yi, and Yukai He. LP-BERT: Multi-task pre-training knowledge graph bert for link prediction. *arXiv preprint arXiv:2201.04843*, abs/2201.04843, 2022. doi:10.48550/arXiv.2201.04843.

105 Michelle M Li, Kexin Huang, and Marinka Zitnik. Graph representation learning in biomedicine and healthcare. *Nature Biomedical Engineering*, 6:1–17, 10 2022. doi:10.1038/s41551-022-00942-x.

106 Qian Li, Daling Wang, Shi Feng, Cheng Niu, and Yifei Zhang. Global graph attention embedding network for relation prediction in knowledge graphs. *IEEE Transactions on Neural Networks and Learning Systems*, 33:6712–6725, 2021. doi:10.1109/tnnls.2021.3083259.

107 Yuezhang Li, Ronghuo Zheng, Tian Tian, Zhiting Hu, Rahul Iyer, and Katia P. Sycara. Joint embedding of hierarchical categories and entities for concept categorization and dataless classification. In *COLING*, volume abs/1607.07956, pages 2678–2688. ACL, 2016. URL: https://aclanthology.org/C16-1252/.

108 Zixuan Li, Xiaolong Jin, Wei Li, Saiping Guan, Jiafeng Guo, Huawei Shen, Yuanzhuo Wang, and Xueqi Cheng. Temporal knowledge graph reasoning based on evolutional representation learning. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*, pages 408–417. ACM, 2021. doi:10.1145/3404835.3462963.

109 Paul Pu Liang, Yiwei Lyu, Xiang Fan, Zetian Wu, Yun Cheng, Jason Wu, Leslie Chen, Peter Wu, Michelle A. Lee, Yuke Zhu, Ruslan Salakhutdinov, and Louis-Philippe Morency. MultiBench: Multiscale benchmarks for multimodal representation learning. *CoRR*, abs/2107.07502, 2021. arXiv:2107.07502, doi:10.48550/arXiv.2107.07502.

110 Siyuan Liao, Shangsong Liang, Zaiqiao Meng, and Qiang Zhang. Learning dynamic embeddings for temporal knowledge graphs. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, pages 535–543. ACM, 2021. doi:10.1145/3437963.3441741.

111 Lifan Lin and Kun She. Tensor decomposition-based temporal knowledge graph embedding. In *2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 969–975. IEEE, IEEE, 11 2020. doi:10.1109/ictai50040.2020.00151.

112 Yankai Lin, Zhiyuan Liu, Huan-Bo Luan, Maosong Sun, Siwei Rao, and Song Liu. Modeling relation paths for representation learning of knowledge bases. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, volume abs/1506.00379, pages 705–714. Association for Computational Linguistics, 2015. doi:10.18653/v1/d15-1082.

113 Yankai Lin, Zhiyuan Liu, Huanbo Luan, Maosong Sun, Siwei Rao, and Song Liu. Modeling relation paths for representation learning of knowledge bases. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, volume abs/1506.00379, pages 705–714. Association for Computational Linguistics, 2015. doi:10.18653/v1/d15-1082.

114 Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. Learning entity and relation embeddings for knowledge graph completion. In *Proceedings of the AAAI conference on artificial intelligence*, volume 29, pages 2181–2187. Association for the Advancement of Artificial Intelligence (AAAI), 2015. doi:10.1609/aaai.v29i1.9491.

115 Sedigheh Mahdavi, Shima Khoshraftar, and Aijun An. dynnode2vec: Scalable dynamic network embedding. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 3762–3765. IEEE, IEEE, 2018. doi:10.1109/bigdata.2018.8621910.

116 Pasquale Minervini, Thomas Demeester, Tim Rocktäschel, and Sebastian Riedel. Adversarial sets for regularising neural link predictors. In *UAI2017, the 33rd Conference on Uncertainty in Artificial Intelligence*, volume abs/1707.07596, pages 1–10. AUAI Press, 2017. URL: http://auai.org/uai2017/proceedings/papers/306.pdf.

117 Seungwhan Moon, Pararth Shah, Anuj Kumar, and Rajen Subba. OpenDialKG: Explainable conversational reasoning with attention-based walks over knowledge graphs. In *Proceedings of the 57th annual meeting of the association for computational linguistics*, pages 845–854. Association for Computational Linguistics, 2019. doi:10.18653/v1/p19-1081.

118 Deepak Nathani, Jatin Chauhan, Charu Sharma, and Manohar Kaul. Learning attention-based embeddings for relation prediction in knowledge graphs. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 4710–4723. Association for Computational Linguistics, 2019. doi:10.18653/V1/P19-1466.

119 Mojtaba Nayyeri, Gokce Muge Cil, Sahar Vahdati, Francesco Osborne, Mahfuzur Rahman, Simone Angioni, Angelo Salatino, Diego Reforgiato Recupero, Nadezhda Vassilyeva, Enrico Motta, et al. Trans4E: Link prediction on scholarly knowledge graphs. *Neurocomputing*, 461:530–542, 10 2021. doi:10.1016/j.neucom.2021.02.100.

120 Mojtaba Nayyeri, Zihao Wang, Mst Akter, Mirza Mohtashim Alam, Md Rashad Al Hasan Rony, Jens Lehmann, Steffen Staab, et al. Integrating knowledge graph embedding and pretrained language models in hypercomplex spaces. *arXiv preprint arXiv:2208.02743*, abs/2208.02743, 2022. doi:10.48550/arxiv.2208.02743.

121 Mojtaba Nayyeri, Bo Xiong, Majid Mohammadi, Mst. Mahfuja Akter, Mirza Mohtashim Alam, Jens Lehmann, and Steffen Staab. Knowledge graph embeddings using neural Itô process: From multiple walks to stochastic trajectories. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 7165–7179, 2023. doi:10.18653/v1/2023.findings-acl.448.

122 Thomas Neumann and Guido Moerkotte. Characteristic sets: Accurate cardinality estimation for RDF queries with multiple joins. In *2011 IEEE 27th International Conference on Data Engineering*, pages 984–994. IEEE, IEEE, 2011. doi:10.1109/icde.2011.5767868.

123 Maximilian Nickel, Lorenzo Rosasco, and Tomaso Poggio. Holographic embeddings of knowledge graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30, pages 1955–1961. Association for the Advancement of Artificial Intelligence (AAAI), 2016. doi:10.1609/aaai.v30i1.10314.

124 Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. A three-way model for collective learning on multi-relational data. In *Proceedings of the International Conference on Machine Learning*, volume 11, pages 3104482–3104584. Omnipress, 2011.

125 Matteo Palmonari and Pasquale Minervini. Knowledge graph embeddings and explainable AI. *Knowledge Graphs for Explainable Artificial Intelligence: Foundations, Applications and Challenges*, 47:49, 2020. doi:10.3233/ssw200011.

126 Enrico Palumbo, Diego Monti, Giuseppe Rizzo, Raphaël Troncy, and Elena Baralis. entity2rec: Property-specific knowledge graph embeddings for item recommendation. *Expert Systems with Applications*, 151:113235, 8 2020. doi:10.1016/j.eswa.2020.113235.

127 Maria Angela Pellegrino, Abdulrahman Altabba, Martina Garofalo, Petar Ristoski, and Michael Cochez. GEval: a modular and extensible evaluation framework for graph embedding techniques. In *European Semantic Web Conference*, volume 12123, pages 565–582. Springer, Springer International Publishing, 2020. doi:10.1007/978-3-030-49461-2_33.

128 Pouya Pezeshkpour, Liyan Chen, and Sameer Singh. Embedding multimodal relational data for knowledge base completion. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, volume abs/1809.01341, pages 3208–3218, Brussels, Belgium, 2018. Association for Computational Linguistics. doi:10.18653/v1/D18-1359.

129 Pouya Pezeshkpour, CA Irvine, Yifan Tian, and Sameer Singh. Investigating robustness and interpretability of link prediction via adversarial modifications. In *Proceedings of NAACL-HLT*, pages 3336–3347. Association for Computational Linguistics, 2019. doi:10.18653/v1/n19-1337.

130 R.J. Pittman, Srivastava. Amit, Sanjika Hewavitharana, Ajinkya Kale, and Saab Mansour. Cracking the code on conversational commerce. ebay blog. 2017. Accessed: 2022-07-03.

131 Jan Portisch, Nicolas Heist, and Heiko Paulheim. Knowledge graph embedding for data mining vs. knowledge graph embedding for link prediction - two sides of the same coin? *Semantic Web*, 13:399–422, 2022. doi:10.3233/sw-212892.

132 Jan Portisch and Heiko Paulheim. The DLCC node classification benchmark for analyzing knowledge graph embeddings. In *The Semantic Web–ISWC 2022: 21st International Semantic Web Conference, Virtual Event, October 23–27, 2022, Pro-*

ceedings, volume abs/2207.06014, pages 592–609. Springer, Springer International Publishing, 2022. doi:10.1007/978-3-031-19433-7_34.

133 Jan Portisch and Heiko Paulheim. Walk this way! entity walks and property walks for RDF2vec. *CoRR*, abs/2204.02777:133–137, 2022. URL: http://arxiv.org/pdf/2204.02777, doi:10.48550/arxiv.2204.02777.

134 Kashif Rabbani, Matteo Lissandrini, and Katja Hose. Extraction of validating shapes from very large knowledge graphs. *Proceedings of the VLDB Endowment*, 16(5):1023–1032, 1 2023. URL: https://www.vldb.org/pvldb/vol16/p1023-rabbani.pdf, doi:10.14778/3579075.3579078.

135 Wessel Radstok, Mel Chekol, and Yannis Velegrakis. Leveraging static models for link prediction in temporal knowledge graphs. In *2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 1034–1041. IEEE, IEEE, 11 2021. URL: http://arxiv.org/pdf/2106.15223, doi:10.1109/ictai52525.2021.00165.

136 Wessel Radstok, Melisachew Wudage Chekol, and Mirko Tobias Schäfer. Are knowledge graph embedding models biased, or is it the data that they are trained on? In *Wikidata@ISWC*, volume 2982. CEUR-WS.org, 2021. URL: https://ceur-ws.org/Vol-2982/paper-5.pdf.

137 Yves Raimond, Tristan Ferne, Michael Smethurst, and Gareth Adams. The BBC world service archive prototype. *Journal of web semantics*, 27:2–9, 8 2014. doi:10.2139/ssrn.3199103.

138 Hongyu Ren, Mikhail Galkin, Michael Cochez, Zhaocheng Zhu, and Jure Leskovec. Neural graph reasoning: Complex logical query answering meets graph databases. *arXiv.org*, abs/2303.14617, 3 2023. URL: http://arxiv.org/pdf/2303.14617, arXiv:2303.14617, doi:10.48550/arxiv.2303.14617.

139 Hongyu Ren, Weihua Hu, and Jure Leskovec. Query2box: Reasoning over knowledge graphs in vector space using box embeddings. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*, volume abs/2002.05969. OpenReview.net, 2 2020. URL: https://openreview.net/forum?id=BJgr4kSFDS.

140 Hongyu Ren and Jure Leskovec. Beta embeddings for multi-hop logical reasoning in knowledge graphs. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, volume abs/2010.11465, 2020. URL: https://proceedings.neurips.cc/paper/2020/hash/e43739bba7cdb577e9e3e4e42447f5a5-Abstract.html.

141 Petar Ristoski and Heiko Paulheim. Rdf2Vec: RDF graph embeddings for data mining. In *Proceedings of the International Semantic Web Conference*, volume 9981, pages 498–514. Springer, Springer International Publishing, 2016. doi:10.1007/978-3-319-46523-4_30.

142 Alexander Rives, Joshua Meier, Tom Sercu, Siddharth Goyal, Zeming Lin, Jason Liu, Demi Guo, Myle Ott, C. Lawrence Zitnick, Jerry Ma, and Rob Fergus. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 118(15):e2016239118, April 2021. doi:10.1073/pnas.2016239118.

143 Tim Rocktäschel, Sameer Singh, and Sebastian Riedel. Injecting logical background knowledge into embeddings for relation extraction. In *Proceedings of the 2015 conference of the north American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1119–1129. Association for Computational Linguistics, 2015. doi:10.3115/v1/n15-1118.

144 Jerret Ross, Brian Belgodere, Vijil Chenthamarakshan, Inkit Padhi, Youssef Mroueh, and Payel Das. Molformer: Large scale chemical language representations capture molecular structure and properties. 5 2022. doi:10.21203/rs.3.rs-1570270/v1.

145 Tara Safavi and Danai Koutra. CoDEx: A comprehensive knowledge graph completion benchmark. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8328–8350. Association for Computational Linguistics, 2020. doi:10.18653/v1/2020.emnlp-main.669.

146 Hooman Peiro Sajjad, Andrew Docherty, and Yuriy Tyshetskiy. Efficient representation learning using random walks for dynamic graphs. *arXiv preprint arXiv:1901.01346*, abs/1901.01346, 1 2019. doi:10.48550/arXiv.1901.01346.

147 Apoorv Saxena, Adrian Kochsiek, and Rainer Gemulla. Sequence-to-sequence knowledge graph completion and question answering. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume abs/2203.10321, pages 2814–2828, Dublin, Ireland, 2022. Association for Computational Linguistics. doi:10.18653/v1/2022.acl-long.201.

148 Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In *Proceedings of the European semantic web conference*, volume 10843, pages 593–607. Springer International Publishing, 2018. doi:10.1007/978-3-319-93417-4_38.

149 Edward W Schneider. Course modularization applied: The interface system and its implications for sequence control and data analysis. 11 1973.

150 Lena Schwertmann, Manoj Prabhakar Kannan Ravi, and Gerard de Melo. Model-agnostic bias measurement in link prediction. In *Findings of the Association for Computational Linguistics: EACL 2023, Dubrovnik, Croatia, May 2-6, 2023*, pages 1587–1603. Association for Computational Linguistics, 2023. doi:10.18653/V1/2023.FINDINGS-EACL.121.

151 Pengpeng Shao, Dawei Zhang, Guohua Yang, Jianhua Tao, Feihu Che, and Tong Liu. Tucker decomposition-based temporal knowledge graph completion. *Knowledge-Based Systems*, 238:107841, 2 2022. URL: https://arxiv.org/pdf/2011.07751, doi:10.1016/j.knosys.2021.107841.

152 Yinghan Shen, Xuhui Jiang, Zijian Li, Yuanzhuo Wang, Chengjin Xu, Huawei Shen, and Xueqi

Cheng. UniSKGRep: A unified representation learning framework of social network and knowledge graph. *Neural Networks*, 158:142–153, 1 2023. doi:10.1016/j.neunet.2022.11.010.

153 Baoxu Shi and Tim Weninger. ProjE: Embedding projection for knowledge graph completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, pages 1236–1242. AAAI Press, 2017. doi:10.1609/aaai.v31i1.10677.

154 Khemraj Shukla, Mengjia Xu, Nathaniel Trask, and George E. Karniadakis. Scalable algorithms for physics-informed neural and graph networks. *Data-Centric Engineering*, 3:e24, 2022. doi:10.1017/dce.2022.24.

155 Amit Singhal. Introducing the knowledge graph: things, not strings. Google Blog. 2012. Accessed: 2022-07-03.

156 Richard Socher, Danqi Chen, Christopher D Manning, and Andrew Ng. Reasoning with neural tensor networks for knowledge base completion. In *Proceedings of the Advances in neural information processing systems*, pages 926–934, 2013.

157 Ran Song, Shizhu He, Shengxiang Gao, Li Cai, Kang Liu, Zhengtao Yu, and Jun Zhao. Multilingual knowledge graph completion from pretrained language models with knowledge constraints. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 7709–7721. Association for Computational Linguistics, 2023. doi:10.18653/v1/2023.findings-acl.488.

158 Fabian M Suchanek, Gjergji Kasneci, and Gerhard Weikum. YAGO: a core of semantic knowledge. In *Proceedings of the 16th international conference on World Wide Web*, pages 697–706. ACM, 2007. doi:10.1145/1242572.1242667.

159 Zequn Sun, Qingheng Zhang, Wei Hu, Chengming Wang, Muhao Chen, Farahnaz Akrami, and Chengkai Li. A benchmarking study of embedding-based entity alignment for knowledge graphs. *Proc. VLDB Endow.*, 13(12):2326–2340, jul 2020. doi:10.14778/3407790.3407828.

160 Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. Rotate: Knowledge graph embedding by relational rotation in complex space. In *Proceedings of the 7th International Conference on Learning Representations*. OpenReview.net, 2 2019. URL: https://openreview.net/forum?id=HkgEQnRqYQ.

161 Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. RotatE: Knowledge graph embedding by relational rotation in complex space. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*, 2019. URL: https://openreview.net/forum?id=HkgEQnRqYQ.

162 Zhu Sun, Jie Yang, Jie Zhang, Alessandro Bozzon, Long-Kai Huang, and Chi Xu. Recurrent knowledge graph embedding for effective recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 297–305. ACM, 2018. doi:10.1145/3240323.3240361.

163 Xiaoli Tang, Rui Yuan, Qianyu Li, Tengyun Wang, Haizhi Yang, Yundong Cai, and Hengjie Song. Timespan-aware dynamic knowledge graph embedding by incorporating temporal evolution. *IEEE Access*, 8:6849–6860, 1 2020. URL: https://ieeexplore.ieee.org/ielx7/6287639/8948470/08950081.pdf, doi:10.1109/access.2020.2964028.

164 Xing Tang, Ling Chen, Jun Cui, and Baogang Wei. Knowledge representation learning with entity descriptions, hierarchical types, and textual relations. *Information Processing & Management*, 56(3):809–822, 5 2019. doi:10.1016/j.ipm.2019.01.005.

165 Yi Tay, Anh Luu, and Siu Cheung Hui. Non-parametric estimation of multiple embeddings for link prediction on dynamic knowledge graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, pages 1243–1249. Association for the Advancement of Artificial Intelligence (AAAI), 2017. doi:10.1609/aaai.v31i1.10685.

166 Yi Tay, Luu Anh Tuan, Minh C Phan, and Siu Cheung Hui. Multi-task neural network for non-discrete attribute prediction in knowledge graphs. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 1029–1038. ACM, 2017. doi:10.1145/3132847.3132937.

167 Komal Teru, Etienne Denis, and Will Hamilton. Inductive relation prediction by subgraph reasoning. In *International Conference on Machine Learning*, volume 119, pages 9448–9457. PMLR, PMLR, 2020. URL: http://proceedings.mlr.press/v119/teru20a.html.

168 Vinh Tong, Dat Quoc Nguyen, Trung Thanh Huynh, Tam Thanh Nguyen, Quoc Viet Hung Nguyen, and Mathias Niepert. Joint multilingual knowledge graph completion and alignment. In *Findings of the Association for Computational Linguistics: EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 4646–4658. Association for Computational Linguistics, 2022. doi:10.18653/v1/2022.findings-emnlp.341.

169 Kristina Toutanova, Danqi Chen, Patrick Pantel, Hoifung Poon, Pallavi Choudhury, and Michael Gamon. Representing text for joint embedding of text and knowledge bases. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1499–1509. Association for Computational Linguistics, 2015. doi:10.18653/v1/D15-1174.

170 Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard. Complex embeddings for simple link prediction. In *International conference on machine learning*, volume 48, pages 2071–2080. PMLR, JMLR.org, 2016.

171 Neil Veira, Brian Keng, Kanchana Padmanabhan, and Andreas G Veneris. Unsupervised embedding enhancements of knowledge graphs using textual associations. In *IJCAI*, pages 5218–5225. International Joint Conferences on Artificial Intelligence Organization, 2019. doi:10.24963/ijcai.2019/725.

172 Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*. OpenReview.net, 2018. URL: https://openreview.net/forum?id=rJXMpikCZ.

173 Denny Vrandečić and Markus Krötzsch. Wikidata: a free collaborative knowledgebase. *Communications of the ACM*, 57:78–85, 2014. doi:10.1145/2629489.

174 Bo Wang, Tao Shen, Guodong Long, Tianyi Zhou, Ying Wang, and Yi Chang. Structure-augmented text representation learning for efficient knowledge graph completion. In *Proceedings of the Web Conference 2021*, pages 1737–1748. ACM, 2021. doi:10.1145/3442381.3450043.

175 Hongwei Wang, Fuzheng Zhang, Jialin Wang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. Ripple-Net: Propagating user preferences on the knowledge graph for recommender systems. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 417–426. ACM, 3 2018. doi:10.1145/3269206.3271739.

176 Hongwei Wang, Fuzheng Zhang, Xing Xie, and Minyi Guo. DKN: Deep knowledge-aware network for news recommendation. In *Proceedings of the 2018 world wide web conference*, pages 1835–1844. ACM, 2018. doi:10.1145/3178876.3186175.

177 Jingbin Wang, Wang Zhang, Xinyuan Chen, Jing Lei, and Xiaolian Lai. 3DRTE: 3d rotation embedding in temporal knowledge graph. *IEEE Access*, 8:207515–207523, 2020. URL: https://doi.org/10.1109/ACCESS.2020.3036897, doi:10.1109/access.2020.3036897.

178 Rui Wang, Bicheng Li, Shengwei Hu, Wenqian Du, and Min Zhang. Knowledge graph embedding via graph attenuated attention networks. *IEEE Access*, 8:5212–5224, 2019. doi:10.1109/access.2019.2963367.

179 Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. KGAT: Knowledge graph attention network for recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 950–958. ACM, 2019. doi:10.1145/3292500.3330989.

180 Xiang Wang, Dingxian Wang, Canran Xu, Xiangnan He, Yixin Cao, and Tat-Seng Chua. Explainable reasoning over knowledge graphs for recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 5329–5336. Association for the Advancement of Artificial Intelligence (AAAI), 7 2019. doi:10.1609/aaai.v33i01.33015329.

181 Xiaozhi Wang, Tianyu Gao, Zhaocheng Zhu, Zhengyan Zhang, Zhiyuan Liu, Juanzi Li, and Jian Tang. KEPLER: A unified model for knowledge embedding and pre-trained language representation. *Transactions of the Association for Computational Linguistics*, 9:176–194, 2021. doi:10.1162/tacl_a_00360.

182 Yashen Wang and Huanhuan Zhang. HARP: A novel hierarchical attention model for relation prediction. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 15:1–22, 2021. doi:10.1145/3424673.

183 Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. Knowledge graph and text jointly embedding. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1591–1601, Doha, Qatar, 2014. Association for Computational Linguistics. doi:10.3115/v1/D14-1167.

184 Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. Knowledge graph embedding by translating on hyperplanes. In *Proceedings of the AAAI conference on artificial intelligence*, volume 28, pages 1112–1119. Association for the Advancement of Artificial Intelligence (AAAI), 2014. doi:10.1609/aaai.v28i1.8870.

185 Zhihao Wang and Xin Li. Hybrid-TE: Hybrid translation-based temporal knowledge graph embedding. In *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 1446–1451. IEEE, IEEE, 2019. doi:10.1109/ictai.2019.00205.

186 Yuyang Wei, Wei Chen, Zhixu Li, and Lei Zhao. Incremental update of knowledge graph embedding by rotating on hyperplanes. In *2021 IEEE International Conference on Web Services (ICWS)*, pages 516–524. IEEE, 2021. doi:10.1109/icws53863.2021.00072.

187 Christopher Wewer, Florian Lemmerich, and Michael Cochez. Updating embeddings for dynamic knowledge graphs. *arXiv.org*, abs/2109.10896, 9 2021. arXiv:2109.10896, doi:10.48550/arXiv.2109.10896.

188 Jiawei Wu, Ruobing Xie, Zhiyuan Liu, and Maosong Sun. Knowledge representation via joint learning of sequential text and knowledge graphs. *arXiv preprint arXiv:1609.07075*, abs/1609.07075, 9 2016. doi:10.48550/arXiv.1609.07075.

189 Tianxing Wu, Arijit Khan, Melvin Yong, Guilin Qi, and Meng Wang. Efficiently embedding dynamic knowledge graphs. *Knowledge-Based Systems*, 250:109124, 8 2022. URL: https://arxiv.org/pdf/1910.06708, doi:10.1016/j.knosys.2022.109124.

190 Yanrong Wu and Zhichun Wang. Knowledge graph embedding with numeric attributes of entities. In *Proceedings of the Rep4NLP@ACL*, pages 132–136. Association for Computational Linguistics, 2018. doi:10.18653/v1/w18-3017.

191 Yikun Xian, Zuohui Fu, S. Muthukrishnan, Gerard de Melo, and Yongfeng Zhang. Reinforcement knowledge graph reasoning for explainable recommendation. In *Proceedings of SIGIR 2019*, pages 285–294, 2019. doi:10.1145/3331184.3331203.

192 Han Xiao, Minlie Huang, Yu Hao, and Xiaoyan Zhu. TransG: A generative mixture model for knowledge graph embedding. *arXiv preprint arXiv:1509.05488*, abs/1509.05488, 2015. doi:10.48550/arXiv.1509.05488.

193 Ruobing Xie, Zhiyuan Liu, Huanbo Luan, and Maosong Sun. Image-embodied knowledge representation learning. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 3140–3146. International Joint Conferences on Artificial Intelligence Organization, 2017. doi:10.24963/ijcai.2017/438.

194 Ruobing Xie, Zhiyuan Liu, Maosong Sun, et al. Representation learning of knowledge graphs with hierarchical types. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 2965–2971. IJCAI/AAAI Press, 2016. URL: http://www.ijcai.org/Abstract/16/421.

195 Chengjin Xu, Yung-Yu Chen, Mojtaba Nayyeri, and Jens Lehmann. Temporal knowledge graph completion using a linear temporal regularizer and multivector embeddings. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2569–2578.

Association for Computational Linguistics, 2021. doi:10.18653/v1/2021.naacl-main.202.

196 Jiacheng Xu, Xipeng Qiu, Kan Chen, and Xuanjing Huang. Knowledge graph representation with jointly structural and textual encoding. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 1318–1324. International Joint Conferences on Artificial Intelligence Organization, 8 2017. doi:10.24963/ijcai.2017/183.

197 Youri Xu, E Haihong, Meina Song, Wenyu Song, Xiaodong Lv, Wang Haotian, and Yang Jinrui. RTFE: A recursive temporal fact embedding framework for temporal knowledge graph completion. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5671–5681. Association for Computational Linguistics, 2021. doi:10.18653/v1/2021.naacl-main.451.

198 Bishan Yang, Wen tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. Embedding entities and relations for learning and inference in knowledge bases. In *Proceedings of the 3rd International Conference on Learning Representations*, 2015.

199 Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. Embedding entities and relations for learning and inference in knowledge bases. *arXiv preprint arXiv:1412.6575*, 12 2014. doi:https://doi.org/10.48550/arXiv.1412.6575.

200 Siyu Yao, Ruijie Wang, Shen Sun, Derui Bu, and Jun Liu. Joint embedding learning of educational knowledge graphs. *Artificial Intelligence Supported Educational Technologies*, pages 209–224, 2020. doi:10.1007/978-3-030-41099-5_12.

201 Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L Hamilton, and Jure Leskovec. Graph convolutional neural networks for web-scale recommender systems. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 974–983. ACM, 2018. doi:10.1145/3219819.3219890.

202 Zhitao Ying, Dylan Bourgeois, Jiaxuan You, Marinka Zitnik, and Jure Leskovec. GNNExplainer: Generating explanations for graph neural networks. *Advances in neural information processing systems*, 32:9240–9251, 2019.

203 Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, Jun Zhao, Zheng Liu, Jing Li, and Maosong Sun. Recursive neural networks for complex relation extraction. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 945–951, 2014.

204 Chuxu Zhang, Huaxiu Yao, Lu Yu, Chao Huang, Dongjin Song, Haifeng Chen, Meng Jiang, and Nitesh V Chawla. Inductive contextual relation learning for personalization. *ACM Transactions on Information Systems (TOIS)*, 39(3):1–22, 7 2021. doi:10.1145/3450353.

205 Fuxiang Zhang, Xin Wang, Zhao Li, and Jianxin Li. TransRHS: a representation learning method for knowledge graphs with relation hierarchical structure. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, pages 2987–2993, 2021.

206 Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma. Collaborative knowledge base embedding for recommender systems. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 353–362. ACM, 2016. doi:10.1145/2939672.2939673.

207 Hanwang Zhang, Zawlin Kyaw, Shih-Fu Chang, and Tat-Seng Chua. Visual translation embedding network for visual relation detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5532–5540. IEEE, 2017. doi:10.1109/cvpr.2017.331.

208 Wen Zhang, Jiaoyan Chen, Juan Li, Zezhong Xu, Jeff Z Pan, and Huajun Chen. Knowledge graph reasoning with logics and embeddings: Survey and perspective. *arXiv preprint arXiv:2202.07412*, abs/2202.07412, 2022. doi:10.48550/arXiv.2202.07412.

209 Wen Zhang, Shumin Deng, Mingyang Chen, Liang Wang, Qiang Chen, Feiyu Xiong, Xiangwen Liu, and Huajun Chen. Knowledge graph embedding in e-commerce applications: Attentive reasoning, explanations, and transferable rules. In *Proceedings of the 10th International Joint Conference on Knowledge Graphs*, pages 71–79. ACM, 2021. doi:10.1145/3502223.3502232.

210 Zhanqiu Zhang, Jianyu Cai, Yongdong Zhang, and Jie Wang. Learning hierarchy-aware knowledge graph embeddings for link prediction. In *AAAI*, volume 34, pages 3065–3072. AAAI Press, 4 2020. doi:10.1609/aaai.v34i03.5701.

211 Zhao Zhang, Fuzhen Zhuang, Meng Qu, Fen Lin, and Qing He. Knowledge graph embedding with hierarchical relation structure. In *EMNLP*, pages 3198–3207. Association for Computational Linguistics, 2018. doi:10.18653/v1/d18-1358.

212 Zhiyuan Zhang, Xiaoqian Liu, Yi Zhang, Qi Su, Xu Sun, and Bin He. Pretrain-KGE: learning knowledge representation from pretrained language models. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 259–266. Association for Computational Linguistics, 2020. doi:10.18653/v1/2020.findings-emnlp.25.

213 Cunchao Zhu, Muhao Chen, Changjun Fan, Guangquan Cheng, and Yan Zhang. Learning from history: Modeling temporal knowledge graphs with sequential copy-generation networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 4732–4740. Association for the Advancement of Artificial Intelligence (AAAI), 5 2021. doi:10.1609/aaai.v35i5.16604.

214 Zhaocheng Zhu, Zuobai Zhang, Louis-Pascal Xhonneux, and Jian Tang. Neural bellman-ford networks: A general graph neural network framework for link prediction. *Advances in Neural Information Processing Systems*, 34:29476–29490, 2021. URL: https://proceedings.neurips.cc/paper/2021/hash/f6a673f09493afcd8b129a0bcf1cd5bc-Abstract.html.

215 Amal Zouaq and Felix Martel. What is the schema of your knowledge graph? leveraging knowledge graph embeddings and clustering for expressive taxonomy learning. In *Proceedings of the international workshop on semantic big data*, pages 1–6. ACM, 2020. doi:10.1145/3391274.3393637.