A Survey of Knowledge Graph Reasoning on Graph Types: Static, Dynamic, and Multimodal

Ke Liang, Lingyuan Meng, Meng Liu, Yue Liu, Wenxuan Tu, Siwei Wang, Sihang Zhou, Xinwang Liu†, *Senior Member, IEEE*, Fuchun Sun, *Fellow, IEEE*

Abstract—Knowledge graph reasoning (KGR), aiming to deduce new facts from existing facts based on mined logic rules underlying knowledge graphs (KGs), has become a fast-growing research direction. It has been proven to significantly benefit the usage of KGs in many AI applications, such as question answering and recommendation systems, etc. According to the graph types, the existing KGR models can be roughly divided into three categories, *i.e.,* static models, temporal models, and multi-modal models. The early works in this domain mainly focus on static KGR and tend to directly apply general knowledge graph embedding models to the reasoning task. However, these models are not suitable for more complex but practical tasks, such as inductive static KGR, temporal KGR, and multi-modal KGR. To this end, multiple works have been developed recently, but no survey papers and open-source repositories comprehensively summarize and discuss models in this important direction. To fill the gap, we conduct a survey for knowledge graph reasoning tracing from static to temporal and then to multi-modal KGs. Concretely, the preliminaries, summaries of KGR models, and typical datasets are introduced and discussed consequently. Moreover, we discuss the challenges and potential opportunities. The corresponding open-source repository is shared on GitHub: https://github.com/LIANGKE23/Awesome-Knowledge-Graph-Reasoning.

Index Terms—Knowledge Graph Reasoning, Knowledge Graph, Temporal Knowledge Graph, Multi-Modal Knowledge Graph.

✦

1 INTRODUCTION

H UMANS learn skills from two main sources, *i.e.,* specialized books and working experiences. For example, a good doctor needs to get knowledge from school and practice experiences from the hospital. However, most existing artificial intelligence (AI) models only imitate the learning procedure from experiences while ignoring the former, thus making them less explainable, as well as sub-optimal performances. To alleviate the problem, many researchers regard the knowledge graphs (KGs), which store the human knowledge facts in intuitive graph structures [1], [2], as potential solutions since they can be easily referred to as books to provide domain-specific knowledge.

To leverage the prior knowledge from KGs, knowledge graph reasoning (KGR), aiming to deduce new facts from existing facts based on the derived underlying logic rules, has drawn increasing attention these years. Specifically, KGR models derive the logic rules *(A, father of, B)*∧*(A, husband of, C)*→*(C, mother of, B)* from existing facts so that the new fact *(Savannah, mother of, Bronny)* in Fig. 1 (a) can be inferred. These KGR models are proven to significantly benefit the usage of KGs in many applications, *e.g.,* question answering [3], recommendation system [4], information extraction [5], image classification [6], etc.

According to the graph types (See Fig. 1), knowledge graph reasoning can be categorized into three categories, *i.e.,* static

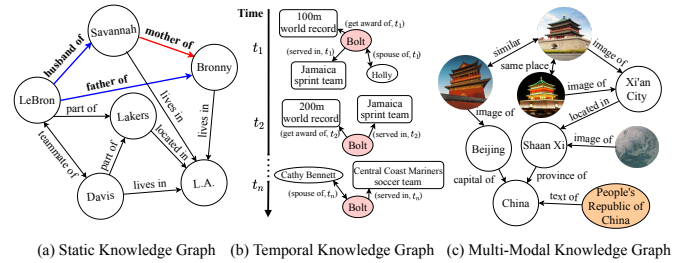(a) Static Knowledge Graph　(b) Temporal Knowledge Graph　(c) Multi-Modal Knowledge Graph

Figure 1: Examples of three categories of the knowledge graph, *i.e.,* static, temporal, and multi-modal knowledge graph.

KGR, temporal KGR, and multi-modal KGR. Specifically, the early works mainly focus on static KGs and tend to directly apply general knowledge graph embedding (KGE) models to the reasoning task, which shows great potential for static KGR in transductive scenarios. However, their expressive abilities are limited for more complex but practical tasks, such as inductive static KGR, temporal KGR, and multi-modal KGR. Note that transductive and inductive are two different reasoning scenarios related to static KGR, which are defined in Sec. 2.2. To address such problems, a large amount of KGR models have been developed recently. For instance, GraIL [7] first performs inductive reasoning based on graph neural networks (GNNs), and many related works are proposed based on it. Moreover, recurrent neural networks (RNNs) are wildly integrated with GNN models for better expression of the time information for temporal KGR [8], [9], [10].

There are several survey papers for the KGR tasks. For instance, [11] first categorizes the KGR as symbolic reasoning and statistical reasoning, while [12] further summarizes the models into three types, *i.e.,* symbolic, neural, and hybrid. After that, [13] and [14] propose more fine-grained categorizations within logic-

- † *Corresponding Author.*
- *Ke Liang, Lingyuan Meng, Meng Liu, Yue Liu, Wenxuan Tu, Siwei Wang, and Xinwang Liu are with the School of Computer, National University of Defense Technology, Changsha, 410073, China. E-mail: liangke200694@126.com, xinwangliu@nudt.edu.cn.*
- *Sihang Zhou is with the College of Intelligence Science and Technology, National University of Defense Technology, Changsha, 410073, China.*
- *Fuchun Sun is with the Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China.*

based and embedding-based KGR models. However, these works only focus on static KGR but omit the recent progress, especially in the fields of temporal and multi-modal KGR. Moreover, none of them discuss the reasoning scenarios of the reviewed KGR models, *i.e.,* transductive, inductive, interpolation, and extrapolation.

To fill the gap, we conduct a survey for knowledge graph reasoning, tracing from static to temporal and then to multi-modal KGs. Concretely, we first briefly introduce the preliminaries. Then, we systematically review the recent KGR models and typical datasets according to the knowledge graph types. Note that the discussion for reasoning scenarios is also presented for the reviewed models. Furthermore, the challenges and potential opportunities are summarized. Our main contributions are shown as follows:

- **Comprehensive Review**. We systematically review the **161** knowledge graph reasoning models, which are **three** times more than others, and comprehensively collect **67** typical datasets according to the graph types, *i.e.,* static, temporal, and multi-modal knowledge graphs. Moreover, we analyze techniques and reasoning scenarios of the reviewed models. To our knowledge, we are the first survey to review the models with the criterion of graph types and reasoning scenarios.
- **Insightful Analysis**. We analyze the strengths and weaknesses of the existing KGR models and their suitable scope, which will provide the readers with useful guidance to select the baselines for their research.
- **Potential Opportunity**. We summarize the challenges of knowledge graph reasoning and point out some potential opportunities which may enlighten the readers.
- **Open-source Resource**. We share the collection of the state-of-the-art KGR models and related datasets on GitHub: https://github.com/LIANGKE23/Awesome-Knowledge-Graph-Reasoning.

The structure of this survey is organized as shown in Fig. 2. Sec. 2 briefly introduces the preliminary. Then, a comprehensive review of recent SOTA KGR models over different KG types is presented in Sec. 3. After that, Sec. 4 lists the typical benchmark KGR datasets. Later, we discuss some challenges and opportunities for KGR in Sec. 5. Finally, Sec. 6 concludes the paper.
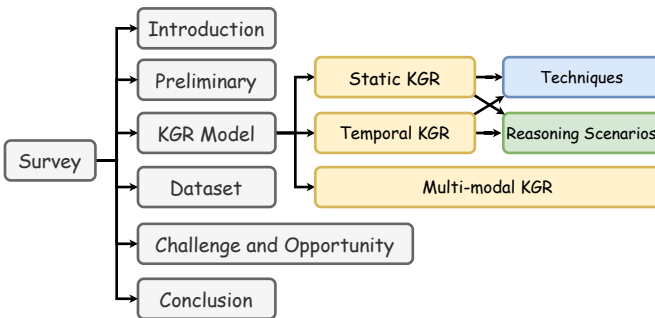

Figure 2: Framework of the survey.

2 PRELIMINARY

This section introduces the preliminaries of knowledge graph reasoning from two aspects, *i.e.,* term definition and problem formulation. We first formally define static, temporal, and multi-modal knowledge graphs. Then, the reasoning tasks over the different knowledge graphs and scenarios are formulated.

Table 1: Notation summary

Notation	Explanation
$\mathcal{SKG}$	Static knowledge graph
$\mathcal{TKG}$	Temporal knowledge graph
$\mathcal{MKG}$	Multi-modal knowledge graph
$\mathcal{E}$	Entity set
$\mathcal{R}$	Relation set
$\mathcal{F}$	The set of facts, *i.e.,* edges
$\mathcal{T}$	The set of the time stamps
$\mathcal{F}_t$	The set of facts at time t
(e_h, r, e_t)	Fact triplet of the head, relation, tail.
(e_h, r, e_t, t)	Fact quadruple of the head, relation, tail, timestamp
(e_h^q, r^q, e_t^q)	Queried fact triplet of head, relation, tail
$\mathbf{e}$	Embedding of entity
$\mathbf{r}$	Embedding of relation
$\mathbf{t}$	Embedding of timestamp

2.1 Definition and Notation

Knowledge graphs (KGs) can be viewed as graphical knowledge bases, thus inheriting most of the functions of traditional knowledge bases [15], such as storing, indexing, etc. Surprisingly, KGs are more suitable for performing reasoning compared to traditional knowledge bases, especially with the development of graph neural networks (GNNs). As shown in Fig. 1, there are three types of knowledge graphs, *i.e.,* static, temporal, and multi-modal knowledge graph. Following previous literature, the formal definition of each type of knowledge graph is declared as follows. Besides, we summarize the notations in Table 1.

Definition 1. Static Knowledge Graph. *Static knowledge graph is defined as* $\mathcal{SKG} = \{\mathcal{E}, \mathcal{R}, \mathcal{F}\}$*, where* $\mathcal{E}$*,* $\mathcal{R}$*, and* $\mathcal{F}$ *represent the sets of entities, relations, and facts. The fact is in a triplet format* $(e_h, r, e_t) \in \mathcal{F}$*, where* $e_h, e_t \in \mathcal{E}$*, and* $r \in \mathcal{R}$ *between them. Note that the static knowledge graph is known as the conventional knowledge graph in [11]. The phrase "static" is used to distinguish it from other types of knowledge graphs.*

Definition 2. Temporal Knowledge Graph. *Temporal knowledge graph is defined as a sequence of static knowledge graphs in different timestamps* $\mathcal{TKG}=\{\mathcal{SKG}_1, \mathcal{SKG}_2, \mathcal{SKG}_3, \cdots, \mathcal{SKG}_t\}$*. The KG at timestamp* t *is defined as* $\mathcal{SKG}_t=\{\mathcal{E}, \mathcal{R}, \mathcal{F}_t\}$*, where* $\mathcal{E}, \mathcal{R}$ *are the sets of entities and relations,* $\mathcal{F}_t$ *is the set of facts at timestamp* $t \in \mathcal{T}$*. The quadruple fact* (e_h, r, e_t, t) *represents that relation* r *exists between head* e_h *and tail* e_t *at timestamp* t*.*

Definition 3. Multi-Modal Knowledge Graph. *Multi-modal knowledge graph* $\mathcal{MKG}$ *is composed of knowledge facts where more than one modalities exist. As an early-stage research field, the relevant definitions are not systematic enough. Generally speaking, according to the representation mode of other modal data, there are two multi-modal KG [16], i.e.,* N-MMKG *and* A-MMKG *(See Fig. 3). Moreover, we also count multi-modal temporal KG as one type of* $\mathcal{MKG}$*, though there are few works.*

2.2 Problem Formulation

Knowledge graph reasoning (KGR), aiming to deduce new facts from existing facts based on the derived underlying logic rules, has drawn increasing attention these years. Specifically, KGR models aim to derive the logic rules *(A, father of, B)*∧*(A, husband of,*
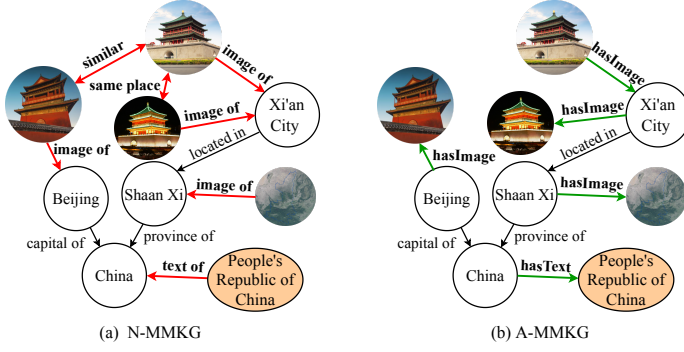
Figure 3: Comparison between two types of multi-modal knowledge graphs. *N-MMKG* represents the multi-modal data as entities, while *A-MMKG* represents multi-modal data as new attributes.

$C) \rightarrow (C, \text{ mother of}, B)$ from existing facts so that the new fact *(Savannah, mother of, Bronny)* in Fig. 1 (a) can be inferred. According to the graph types defined in Sec. 2.1, KGR can be categorized into three types, *i.e.,* static KGR, temporal KGR, and multi-modal KGR. More concretely, given the **queried** facts, the KGR models calculate the likelihood of the queried relation r_q between the queried tail e_t^q and the queried tail e_h^q based on the scoring functions. However, compared to static KGs, the additional information is contained in temporal and multi-modal KGs, such as time information and visual information, thus leading to slight variances of reasoning over different graph types.

This section will formulate the related KGR problems according to the graph types. Moreover, we further illustrate two groups of terms of the reasoning scenarios, *i.e.,* transductive & inductive scenarios and interpolation & extrapolation scenarios.

Static Knowledge Graph Reasoning. Given a static knowledge graph $\mathcal{SKG} = \{\mathcal{E}, \mathcal{R}, \mathcal{F}\}$, knowledge graph reasoning aims to exploit the existing facts to infer queried fact (e_h^q, r_q, e_t^q). Three sub-tasks are defined, *i.e.,* head reasoning $(?, r^q, e_t^q)$, tail inferring $(e_h^q, r^q, ?)$, and relation inferring $(e_h^q, ?, e_t^q)$.

Temporal Knowledge Graph Reasoning. Given a temporal knowledge graph $\mathcal{TKG} = \{\mathcal{SKG}_1, \mathcal{SKG}_2, \mathcal{SKG}_3, \cdots, \mathcal{SKG}_t\}$, where $\mathcal{SKG}_t = \{\mathcal{E}, \mathcal{R}, \mathcal{F}_t\}$ and timestamp $t \in \mathcal{T}$. With the queried fact (e_h^q, r^q, e_t^q, t^q), reasoning can be classified into three types, including entity reasoning, *i.e.,* $(?, r^q, e_t^q, t^q)$ and $(e_h^q, r^q, ?, t^q)$, relation reasoning $(e_h^q, ?, e_t^q, t^q)$, and timestamp reasoning $(e_h^q, r^q, e_t^q, ?)$. The former two sub-tasks are similar to the reasoning on static and multi-modal KGs, while the latter one is unique in temporal KGR.

Multi-modal Knowledge Graph Reasoning. Reasoning over multi-modal knowledge graph reasoning is similar to the reasoning task for the other two KG types, *i.e.,* inferring the missing facts either in triplet or quadruple format. But, since the entities are in more than one modalities in multi-modal KGs, multi-modal KGR generally requires extra knowledge fusion modules for different modalities before fact inference.

Transductive and Inductive Reasoning. According to the visibility of the queried entities and relations in the training procedure, there are two types of reasoning scenarios as shown in Fig. 4, *i.e.,* transductive reasoning, and inductive reasoning. Within transductive reasoning, the entities and relationships in the queried
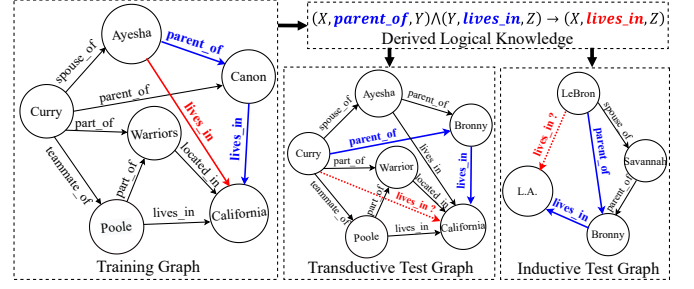

Figure 4: Illustration of transductive and inductive reasoning. In the transductive scenario, entities in test graphs are all seen during the training procedure. While as for the inductive scenario, unseen entities may exist in test graphs.
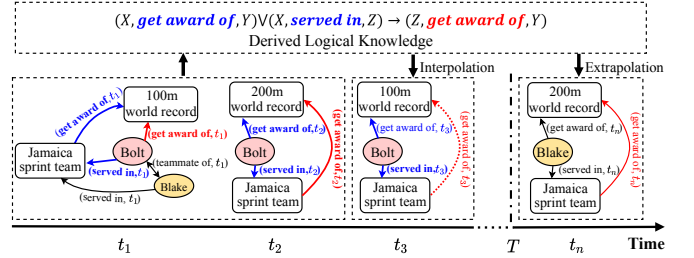

Figure 5: Illustration of interpolation and extrapolation reasoning. The timestamp t for knowledge graph reasoning in the interpolation scenario is seen in the past $(0 \leq t \leq T)$. While the queried facts in the future $(t \geq T)$ for the extrapolation scenario.

fact are all seen in the given knowledge graph, *i.e.,* $e_h^q, e_t^q \in \mathcal{E}$ and $r^q \in \mathcal{R}$. As for inductive reasoning, the candidates for e_h^q, e_t^q and r^q may be beyond the given graph. Note that these two scenarios are mostly discussed in static KGR models.

Interpolation and Extrapolation Reasoning. According to the occurrence time of the queried fact t^q, the reasoning can be divided into two categories as shown in Fig. 5, *i.e.,* interpolation reasoning and extrapolation reasoning. Concretely, given a temporal KG with the timestamps ranging from time 0 to time T, the interpolation reasoning aims to infer the queried facts for time t, where $0 \leq t \leq T$; besides, the extrapolation reasoning aims to infer the queried facts for time t, where $t \geq T$. Note that these two scenarios are mostly discussed in temporal KGR models.

3 KNOWLEDGE GRAPH REASONING MODEL

We comprehensively review **161** knowledge graph reasoning (KGR) models. Specifically, the KGR models are summarized based on graph types, *i.e.,* static, temporal, and multi-modal knowledge graphs.

3.1 Static KGR Model

We first systematically introduce **88** static KGR models according to the techniques. Then, we categorize the reviewed models according to the reasoning scenarios. The taxonomy for static KGR models is shown in Fig. 6.

3.1.1 Embedding-based Model

Embedding-based models learn the embedding vectors based on existing fact triplets and then rank the top k candidate facts
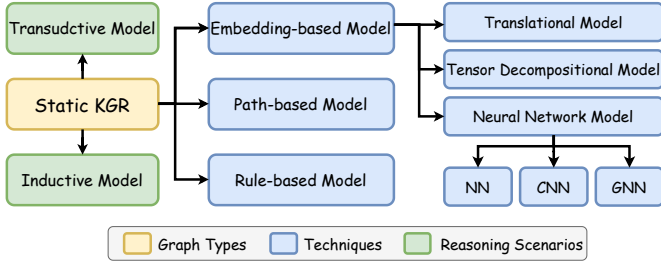
Figure 6: Taxonomy of the static KGR models.

according to the likelihood scores calculated by the scoring functions. There are three types in a majority quantity, *i.e.,* translational models, tensor decompositional models, and neural network models. According to our observation, embedding-based models are much more than other models. Thus, the timeline of the models is presented in Fig. 7 for clear presentation.

Translational Model. Translational models usually regard the relation r as translational transformation to project the entity embedding **e** into the latent space. The distance functions, *e.g.,* *L1-norm* and *L2-norm*, are leveraged for scoring the fact triplets.

TransE [17], as the first translational model, regards the relation as a simple translation operation, $e_h + r \approx e_t$. Although proven effective, it cannot handle some specific relations, such as one-to-many, many-to-one, symmetric and transitive relations. To address these limitations, lots of translational models are developed. The entities are encoded into the relation-specific hyperplane in TransH [18], which achieves better reasoning performance on one-to-many, many-to-one relations. Besides, TransR [19] leverages distinct latent spaces for entities and relations and gets better expressive ability on transitive relation reasoning. Moreover, TransD [20] first considers the scalability issues by leveraging independent projection vectors for entities and relations for large KGs. Afterward, probabilistic principles are integrated to model the uncertainty in KGR. For example, KG2E [21] leverages the Gaussian distribution covariance, and TransG [22] makes use of the Bayesian technique for one-to-many relational facts. Meanwhile, to alleviate heterogeneity and imbalance issues, TranSparse [23] provides an efficient solution by designing adaptive transfer sparser matrices, thus leading to better expressive ability. After that, TorusE [24] projects embeddings in a compact Lie group torus, and MuRP [25] designs a Möbius matrix-vector multiplication and Möbius addition for entity embedding projection, which all show better accuracy and scalability. Meanwhile, TransW [26] first has enriched the entity and relation embedding with the word embeddings, which achieves better performances on inferring facts with unseen entities or relations. Moreover, RotatE [27] proposes a rotation-based translational method with complex-valued embeddings to better infer the symmetry, antisymmetry, inversion, and composition facts. Besides, HAKE [28] models the semantic hierarchy rather than relation patterns based on the polar coordinate space. Then, TransRHS [29] first considers the Relation Hierarchical Structure (RHS) by incorporating RHS seamlessly into the embeddings. Besides, to handle the complex relational facts with a unified model, PairRE [30] models each relation representation with paired vectors to adaptive adjustment for complex relations, and HousE [31] involves a novel parameterization based on the designed Householder transformations for rotation and projection. Nowadays, there are also some interesting attempts for translational models for more sufficient interactions, such as TripleRE [32] and InterHT [33]. TripleRE creatively divides the relationship vector into three parts, takes advantage of the concept of residual, and achieves better performance. While, InterHT enhances the information interactions between the tail and head, which can improve the model capacity.

Tensor Decompositional Model. Tensor decompositional models represent the KGs as three-way tensors, decomposed into a combination of low-dimensional vectors for entities and relations.

RESCAL [34], as the first tensor decompositional model, captures to latent semantics of each entity with vectors and further leverages the matrix to model the pairwise interactions among latent factors as a matrix. However, the model is complex with $O(d^2)$ parameters. To simplify it, DistMult [35] uses bi-linear diagonal matrices to reduce parameters to $O(d)$ per relation. Then, ComplEx [36] generalizes DistMult by using complex-valued embeddings, which improve asymmetric relations modeling. Meanwhile, HolE [37] models the holographic reduced representations and circular correlation, and Analogy [38] designs the bi-linear scoring function with analogical structural constraints for analogical reasoning, which both try to capture rich interactions between entities. Then, some models start to substitute the decompositional operations. SimplE [39] enhances the Canonical Polyadic (CP) decompositial for two independent entity embeddings, and Tucker Decomposition is first used by Tucker [40]. Meanwhile, CrossE [41] considers crossover interactions between entities via a relation-specific interaction matrix. QuatE [42] empowers the semantic matching between head and tail based on relational rotation quaternion representations. Inspired by it, DualE [43] projects the embeddings in dual quaternion space to achieve a unified framework for both translation and rotation operations. Besides, HopfE [44] makes use of both structural and semantic attributes in 4D hyper-sphere space without losing interpretability. Besides expressive ability, efficiency also draws increasing attention these years. A factorized bi-linear pooling model is proposed based on Tucker decomposition, termed LowFER [45], which is more efficient and lightweight. Moreover, QuatRE [46] learns the quaternion embeddings with two novel operations, *i.e.,* enhancing the correlations between entities with hamilton product for entity embeddings and reducing computation by simplifying the translation matrices.

Neural Network Model. Neural network models have yielded remarkable performance for KG reasoning these years due to the expressive ability for representation learning on KGs. Three types are defined based on the NN techniques, *i.e.,* traditional neural network (NN) models, convolutional neural network (CNN) models, and graph neural network (GNN) models.

1) Traditional NN Model: SME [47] first encodes the entities and relations into the latent space using neural networks. Meanwhile, neural tensor networks are used in NTN [48] for relation reasoning in KGs. Then, NAM [49] proposes the relational-modulated neural network (RMNN) and NN model shares variables in ProjE [50], which jointly learns embeddings of the entities and relations via the standard loss function. These traditional NN models show great potential on static KGR while they suffer from learning shallow and less expressive features.

2) CNN Model: To learn deeper features, convolutional neural networks (CNNs) are integrated with KGR models. ConvE [51] first leverages 2D convolutional layers for knowledge graph
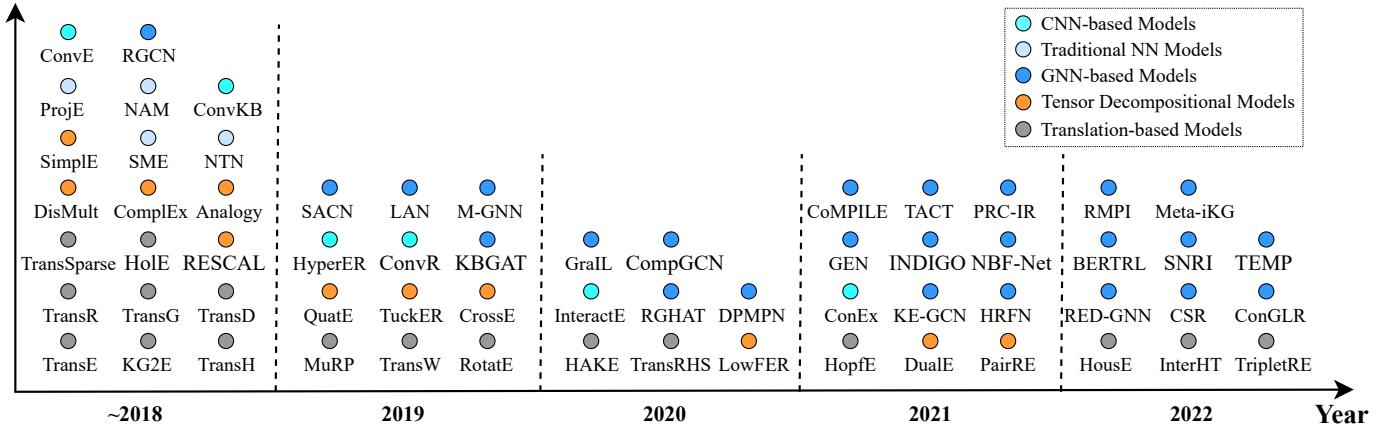
Figure 7: Timeline of the embedding-based models for static KGR.

reasoning. ConvKB [52] extends ConvE by removing the reshaping operation and captures global and transitional characteristics within facts for more informative expression. Later on, HypER [53] makes use of the fully connected layer and relation-specific convolutional filters for better performance. Besides, ConvR [54] designs an adaptive convolutional network designed to maximize entity-relation interactions by constructing convolution filters across entity and relation representations. Moreover, novel operations, *i.e.,* feature reshaping, feature permutation, and circular convolution, are designed in InteractE [55] to handle more interactions within KGs. Meanwhile, ConEx [56] integrates the affine transformation and a Hermitian inner product on complex-valued embeddings with the convolutional operation, which shows good expressiveness. CNN KGR models generally perform better than traditional NN models. However, the information underlying the graph structures cannot be well learned.

3) GNN Model: Graph neural networks, which are widely used for graph tasks, are also rapidly applied to KG reasoning. RGCN [57] uses the relation-specific transformation to aggregate neighborhood information. Then, each entity is encoded into a vector, and the decoder *i.e.,* scoring function, reconstructs the facts based on entity representations. While RGCN omits the variances of entities, which hinders expressive ability. To alleviate it, attention mechanisms are integrated into lots of models, such as M-GNN [58], KBGAT [59] and [60]. In particular, KBGAT [59] leverages attention-based feature embeddings for better reasoning performance. Meanwhile, SACN [61] leverages the weighted graph convolutional network (WGCN) as the encoder and a convolutional network called Conv-TransE as the decoder, which is effective. Afterward, TransGCN [62] trains both relation and entity embeddings simultaneously with the transformation operator for relations. Later on, DPMPN [63] and RGHAT [64] designs a two-GNN framework to simultaneously encode information in different levels separately, *i.e.,* global & local information for DPMPN and relation & entity information for RGHAT. After that, KE-GCN [65] jointly propagates and updates the embedding of both entities and edges. Similarly, COMPGCN [66] also jointly learns the representations with various entity-relation composition operations. Recently, more and more researchers have tried to handle out-of-knowledge-graph scenarios. GEN [67] and HRFN [68] learn entity embeddings based on meta-learning for both seen-to-unseen and unseen-to-unseen facts. INDIGO [69] is then proposed based on a GNN using pair-wise encoding. Besides, the GraIL-based model

is a group of typical GNN models for inductive scenarios. The prototype GraIL [7], as the landmark GNN-based model, first leverages RGCN to perform the reasoning based on the local enclosing subgraph. Based on it, many incremental works are developed, including TACT [70], CoMPILE [71], Meta-iKG [72], SNRI [73], RPC-IR [74], and etc. These GraIL-based models all achieve promising inductive performances. Among them, TACT [70] and CoMPILE [71] both raise the importance of relation embeddings in the task. Concretely, TACT [70] uses topology-aware correlations between relations to generate representations for triplet scoring, which also inspires RMPI [75] and TEMP [76]. Besides, CoMPILE enhances the message interactions between relations and entities with a novel mechanism. After that, motivated by the great success of contrastive mechanisms [77], [78], [79], [80], contrastive learning models have been increasingly proposed, *e.g.,* RPC-IR [74], SNRI [73] etc. Besides, Meta-iKG [72] verifies the effectiveness of meta-learning in the KGR task. After that, researchers try to make the reasoning more efficient. NBF-net [81] and RED-GNN [82] achieve better efficiency by leveraging the traditional algorithm *i.e.,* bellman-ford algorithm and dynamic programming to optimize the propagation strategy in previous GNN models. Besides, pGAT [83] leverages the EM algorithm for efficient learning. Moreover, BERTRL [84] and ConGLR [85] integrates the context for each entity to enhance the reasoning on the knowledge graph. In particular, BERTRL [84] can handle unseen relational facts. Regarding it, CSR [86] deeply mines the logic rules underlying the structure patterns instead of the paths.

3.1.2 Path-based Model

The logical knowledge apparently exists underlying the path between the queried head and tail. According to it, path-based models are proposed to mine information in such paths.

Random walk [111], [112], [113], [114] inferences have been widely investigated in this type of model. For instance, the Path-Ranking Algorithm (PRA) [110] derives the path-based logic rules under path constraints. ProPPR [108] further introduces space similarity heuristics by incorporating textual content to alleviate the feature sparsity issue in PRA. Meanwhile, Neural multi-hop path-based models are also studied for better expressive ability. By iteratively using compositionality, RNNPRA [107] leverages RNN to compose the implications of relational paths for reasoning. LogSumExp [104] designs a logical composition method across all the elements with attention mechanisms for multiple reasoning.

Table 2: Summary of the static knowledge graph reasoning models.

Year	Model	Reasoning Scenrio	Technique	Year	Model	Reasoning Scenrio	Technique
2022	RED-GNN [82]	Inductive	GNN	2019	SACN [61]	Transductive	GNN
2022	ConGLR [85]	Inductive	GNN	2019	KBGAT [59]	Transductive	GNN
2022	TripleRE [32]	Transductive	Translational	2019	LAN [60]	Inductive	GNN
2022	InterHT [33]	Transductive	Translational	2019	CPL [87]	Transductive	Relation Path
2022	HousE [31]	Transductive	Translational	2019	IterE [88]	Inductive	Logic Rule
2022	BERTRL [84]	Inductive	GNN	2019	pLogicNet [89]	Inductive	Logic Rule
2022	SNRI [73]	Inductive	GNN	2019	DRUM [90]	Inductive	Logic Rule
2022	TEMP [76]	Inductive	GNN	2019	RLvLR [91]	Inductive	Logic Rule
2022	RMPI [75]	Inductive	GNN	2019	Neural-Num-LP [92]	Inductive	Logic Rule
2022	Meta-iKG [72]	Inductive	GNN	2018	SimplE [39]	Transductive	Tensor Decompositional
2022	CSR [86]	Inductive	GNN	2018	ConvKB [52]	Transductive	CNN
2022	CURL [93]	Transductive	Relation Path	2018	ConvE [51]	Transductive	CNN
2022	GCR [94]	Inductive	Logic Rule	2018	RGCN [57]	Transductive	GNN
2021	PairRE [30]	Transductive	Translational	2018	M-walk [95]	Transductive	Relation Path
2021	HopfE [44]	Transductive	Tensor Decompositional	2018	MultiHop [96]	Transductive	Relation Path
2021	DualE [43]	Transductive	Tensor Decompositional	2018	DIVA [97]	Transductive	Logic Rule
2021	ConEx [56]	Transductive	CNN	2018	RuleN [98]	Inductive	Logic Rule
2021	KE-GCN [65]	Transductive	GNN	2018	RUGE [99]	Inductive	Logic Rule
2021	HRFN [68]	Inductive	GNN	2017	ANALOGY [38]	Transductive	Tensor Decompositional
2021	GEN [67]	Inductive	GNN	2017	ProjE [50]	Transductive	Traditional NN
2021	INDIGO [69]	Inductive	GNN	2017	MINERVA [100]	Transductive	Relation Path
2021	NBF-Net [81]	Inductive	GNN	2017	DeepPath [101]	Transductive	Relation Path
2021	CoMPILE [71]	Inductive	GNN	2017	NTP [102]	Inductive	Logic Rule
2021	TACT [70]	Inductive	GNN	2017	NeuralLP [103]	Inductive	Logic Rule
2021	RPC-IR [74]	Inductive	GNN	2016	TranSparse [23]	Transductive	Translational
2020	HAKE [28]	Transductive	Translational	2016	TransG [22]	Transductive	Translational
2020	TransRHS [29]	Transductive	Translational	2016	HolE [37]	Transductive	Tensor Decompositional
2020	LowFER [45]	Transductive	Tensor Decompositional	2016	ComplEx [36]	Transductive	Tensor Decompositional
2020	InteractE [55]	Transductive	CNN	2016	NAM [49]	Transductive	Traditional NN
2020	DPMPN [63]	Transductive	GNN	2016	LogSumExp [104]	Transductive	Relation Path
2020	RGHAT [64]	Transductive	GNN	2016	KALE [105]	Inductive	Logic Rule
2020	COMPGCN [66]	Transductive	GNN	2015	TransD [20]	Transductive	Translational
2020	GraIL [7]	Inductive	GNN	2015	TransR [19]	Transductive	Translational
2020	ExpressGNN [106]	Inductive	Logic Rule	2015	KG2E [21]	Transductive	Translational
2020	pGAT [83]	Inductive	GNN	2015	DISTMULT [35]	Transductive	Tensor Decompositional
2019	RotatE [27]	Transductive	Translational	2015	RNNPRA [107]	Transductive	Relation Path
2019	TransW [26]	Inductive	Translational	2014	TransH [18]	Transductive	Translational
2019	MuRP [25]	Transductive	Translational	2014	ProPPR [108]	Transductive	Relation Path
2019	QuatE [42]	Transductive	Tensor Decompositional	2013	AMIE [109]	Inductive	Logic Rule
2019	TuckER [40]	Transductive	Tensor Decompositional	2013	SME [47]	Transductive	Traditional NN
2019	CrossE [41]	Transductive	Tensor Decompositional	2013	NTN [48]	Transductive	Traditional NN
2019	ConvR [54]	Transductive	CNN	2013	TransE [17]	Transductive	Translational
2019	HypER [53]	Transductive	CNN	2011	RESCAL [34]	Transductive	Tensor Decompositional
2019	M-GNN [58]	Transductive	GNN	2010	PRA [110]	Transductive	Relation Path

Then, a unified variational inference framework is proposed by DIVA [97], which separates multi-hop reasoning into two steps, *i.e.,* path-finding and path-reasoning. Deep reinforcement learning (DRL) techniques, such as the Markov decision process (MDP), have recently been used to reformulate path-finding between entities as a sequential decision-making task. The designed reinforcement learning agent learns to find the reasoning paths according to entity interactions, and the corresponding policy gradient is utilized for training. Concretely, different fine-grained manners are employed by different models. For example, DeepPath [101] applies DRL for relational path learning via the novel rewards and relational action spaces, which improves both performance and

efficiency of the models. Meanwhile, MINERVA [100] takes path finding between entities as a sequential optimization problem by maximizing the expected reward [3], which excludes the target answer entity for more capable reasoning. After that, MultiHop [96] designs a soft reward mechanism instead of only relying on binary rewards, as well as the dropout action, which enables more effective path exploration. Besides, Monte Carlo Tree Search (MCTS) is used by M-Walk [95] to generate the path, and CPL [87] proposes collaborative policy learning for path-finding and fact extraction by leveraging the text corpus corresponding to the entities. Moreover, two agents in different levels, *i.e.,* DWARF AGENT at the entity level and GIANT AGENT at the cluster

level, are proposed in CURL [93], which collaborate to achieve optimal reasoning performance.

3.1.3 Rule-based Model

A logic rule is generally defined in the form of B → A, where A is a fact, and B can be a set of facts. To make better use of such symbolic features, Rule-based models are proposed [3].

Logical rules can be extracted from KG for reasoning by rule mining tools, *e.g.,* AMIE [109], RuleN [98], etc. Then, a more scalable rule mining approach via the techniques of rule searching and pruning is designed by RLvLR [91]. After that, how to inject logical rules into embeddings for better reasoning performance has drawn increasing research attention [3]. In general, there are two ways for it, *i.e.,* joint learning and iterative training. For instance, KALE [105] is a unified joint model by leveraging the t-norm fuzzy logical connectives between compatible facts and rule embedding. Besides, RUGE [99] is an iterative model utilizing the soft rules for embedding rectification. Inspired by it, the iterative training strategy composed of embedding learning, axiom induction, and axiom injection is designed by IterE [88]. After that, Researchers tend to integrate neural network techniques into the rule-based models to alleviate the issues of limited expressive ability and huge space consumption of the previous rule-based models. Neural Theorem Provers (NTP) [102] mines logical rules with the designed radial kernel. Besides, NeuralLP [103] leverages attention mechanisms and auxiliary memory to optimize the gradients for mining the rules, and Neural-Num-LP [92] further integrates the cumulative sum operations and dynamic programming with NeuralLP to learn numerical rules. Meanwhile, an end-to-end differentiable rule-based model is proposed in DRUM [90]. Then, the probabilistic logic neural network is designed in pLogicNet [89], which shows great performances for first-order logic mining. Based on it, ExpressGNN [106] further generalizes it by finetuning GNN models for more efficient reasoning. Moreover, GCR [94] achieves promising performance for both reasoning and recommendation by mining the neighborhood information around the queried facts.

3.1.4 Discussion on Static KGR Models

Tab. 2 systematically presents a classification of the static KGR models, especially for the reasoning scenarios, *i.e.,* transduction and induction scenarios. In summary, there are **56** transductive models including TripleRE, InterHT, HouseE, CURL, PairRE, HopfE, DualE, ConEx, KE-GCN, HAKE, TransRHS, LowFER, InteractE, DPMPN, RGHAT, COMPGCN, RotatE, MuRP, QuatE, TuckER, CrossE, ConvR, HypER, M-GNN, SACN, KBGAT, CPL, SimplE, ConvKB, ConvE, RGCN, M-walk, MultiHop, DIVA, ANALOGY, ProjE, MINERVA, DeepPath, TranSparse, TransG, HolE, ComplEx, NAM, LogSumExp, TransD, TransR, KG2E, DISTMULT, RNNPRA, TransH, ProPPR, TransE, SME, NTN, RESCAL, PRA. While there are **32** inductive models, including RED-GNN, ConGLR, BERTRL, SNRI, TEMP, RMPI, Meta-iKG, CSR, GCR, HRFN, GEN, INDIGO, NBF-Net, CoM-PILE, TACT, RPC-IR, GraIL, ExpressGNN, pGAT, TransW, LAN, IterE, pLogicNet, DRUM, RLvLR, Neural-Num-LP, RuleN, RUGE, NTP, NeuralLP, KALE, AMIE.

According to it, we can further get the following observations, which may indicate the scope for different static KGR models and reveal the future trend in static KGR. (1) Embedding-based models generally have better expressive ability but lack explainability. Meanwhile, more attention is currently focused on

developing GNN-based models, which is also shown in Fig. 7. More concretely, with the development of graph neural networks, GNN-based models have shown great potential in various graph-based tasks. The Knowledge graph reasoning models require a high-quality representation of the relational facts and the graph structure, which is most suitable for using GNN. (2) Path-based and Rule-based models are more explainable than embedding-based models, but they usually suffer from limited expressive ability and huge complexity of time and space. (3) Most of the path-based models are more suitable for transductive reasoning due to the path-searching schemes, while Rule-based models naturally inherit the inductive ability due to the generalization of the rule paradigms. (4) For a long time, transductive reasoning models have kept appearing and greatly impacted both academic research and industrial applications. However, due to the issues of scalability and expressive ability issues, researchers have recently focused on developing inductive reasoning models.

3.2 Temporal KGR Model

We first systematically introduce **50** temporal KGR models according to how they integrate time information. Then, we categorize the reviewed models according to the reasoning scenarios. The taxonomy for temporal KGR models is shown in Fig. 8.
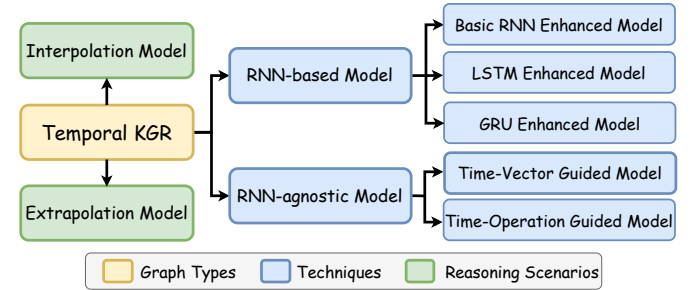

Figure 8: Taxonomy of the temporal KGR models.

3.2.1 RNN-based Model

Recurrent Neural Networks (RNNs) are suitable for mining the changes over time. Thereby many temporal KGR models use RNNs to directly model the temporal information, termed RNN-based models. According to different variants of RNN, the models can be divided into three types, *i.e.,* basic RNN enhanced models, LSTM enhanced models, and GRU enhanced models.

Basic RNN Enhanced Models. Some temporal KGR models can effectively model temporal information using basic RNN models. To name a few, Know-Evolve [163] is a classical temporal KGR model that generates non-linearly entity embeddings over time. RE-NET [151] applies the GCN and RNN models to capture the evolutional dynamics in temporal knowledge graphs to the query over time. EvoKG [150] introduces the RNN model to mine the dynamic evolving structural information and models entity interactions by combining the neighborhood information.

LSTM Enhanced Models. Long Short-Term Memory (LSTM) network is also widely used to mine temporal features in temporal KGR models. For instance, TTransE [155] extends TransE by adding the temporal constraints and encodes time information as translations similar to relationships with an RNN so that these translations move the header representation in the embedded

Table 3: Summary of the temporal knowledge graph reasoning models.

Year	Model	Reasoning Scenrio	Technique	Year	Model	Reasoning Scenrio	Technique
2022	CENET [115]	Extrapolation	Time-Operation	2021	xERTE [116]	Extrapolation	Time-Vector
2022	DA-Net [117]	Extrapolation	Time-Operation	2021	CyGNet [118]	Extrapolation	Time-Vector
2022	HiSMatch [119]	Extrapolation	GRU	2021	TIE [120]	Interpolation	Time-Operation
2022	MetaTKGR [121]	Extrapolation	Time-Operation	2021	TeLM [122]	Interpolation	Time-Operation
2022	rGalT [123]	Extrapolation	Time-Operation	2021	ChronoR [124]	Interpolation	Time-Vector
2022	FILT [125]	Interpolation	Time-Operation	2021	RE-GCN [8]	Extrapolation	GRU
2022	TKGC-AGP [126]	Interpolation	Time-Operation	2021	RTFE [127]	Interpolation	Time-Operation
2022	Tlogic [128]	Extrapolation	Time-Operation	2021	HIP [129]	Extrapolation	GRU
2022	TLT-KGE [130]	Interpolation	Time-Vector	2021	Tpath [131]	Interpolation	LSTM
2022	CEN [132]	Extrapolation	Time-Operation	2020	DyERNIE [133]	Interpolation	Time-Operation
2022	BoxTE [134]	Interpolation	Time-Vector	2020	DacKGR [135]	Interpolation	RNN
2022	TempoQR [136]	Interpolation	Time-Vector	2020	TNTComplEx [137]	Interpolation	Time-Vector
2022	TuckERTNT [138]	Interpolation	Time-Vector	2020	TDGNN [139]	Extrapolation	Time-Operation
2022	DKGE [140]	Interpolation	Time-Operation	2020	ATiSE [141]	Interpolation	Time-Operation
2022	TiRGN [142]	Extrapolation	GRU	2020	Diachronic [143]	Interpolation	Time-Operation
2022	RotateQVS [144]	Interpolation	Time-Vector	2020	TeRo [145]	Interpolation	Time-Operation
2022	ExKGR [146]	Interpolation	LSTM	2020	EvolveGCN [147]	Extrapolation	LSTM+GRU
2022	TRHyTE [148]	Interpolation	GRU	2020	TeMP [149]	Interpolation	GRU
2022	EvoKG [150]	Extrapolation	RNN	2020	RE-NET [151]	Extrapolation	RNN
2021	TPmod [152]	Interpolation	GRU	2019	DyRep [153]	Extrapolation	Time-Operation
2021	TimeTraveler [154]	Extrapolation	LSTM	2018	TTransE [155]	Interpolation	LSTM
2021	CluSTeR [156]	Extrapolation	LSTM+GRU	2018	HyTE [157]	Interpolation	Time-Operation
2021	TPRec [158]	Extrapolation	Time-Operation	2018	ChronoTranslate [159]	Interpolation	Time-Operation
2021	DBKGE [160]	Interpolation	Time-Vector	2018	TA-DISTMULT [161]	Interpolation	LSTM
2021	T-GAP [162]	Interpolation	Time-Vector	2017	Know-Evolve [163]	Extrapolation	RNN

space. TA-TransE and TA-DistMult [161] are also two respectively extended versions of TransE and DistMult that incorporate the temporal embeddings. Furthermore, EvolveGCN [147] adopts the graph convolutional networks (GCNs) to model the graph structure in each static snapshot and utilizes the LSTM model (also can utilize GRU model) to evolve the GCN parameters over time. CluSTeR [156] adopts reinforcement learning to discover evolutional patterns with both LSTM and GRU models in Temporal KGs over time. DacKGR [135] performs multi-hop path-based reasoning on sparse temporal KGs by using time information for dynamic prediction. To capture the timespan information and guide the model learning, TimeTraveler [154] proposes a novel relative time encoding module and a time-shaped reward module based on Dirichlet distribution. TPath [131] also introduces LSTM model to mine the current environment information and then generates relation embeddings and temporal embeddings to the environment through activation functions. ExKGR [146] introduces LSTM for entity reasoning in temporal KGs and provides the reasoning paths.

GRU Enhanced Models. GRU-based models have got a lot of attention these years. More recently, TeMP [149] is proposed, which leverages message-passing graph neural networks (MPNNs) to learn structure-based entity representations at each timestamp, and then combines representations from all timestamps using an encoder. RE-GCN [8] focuses on the evolutional dynamics in temporal KGs and generates entity embeddings by modeling the KG sequence of a fixed length at the latest a few timestamps. TPmod [152] aggregates the attributes of entities and relations and

learns dynamic weights to different events. HIP network [129] passes information from temporal, structural, and repetitive perspectives, which are used to mine the graph's dynamic evolution, the interactions of events at the same time step, and the known events respectively. TRHyTE [148] uses GRU to first transform entities into latent space and then encode facts into temporal-relational hyper-planes for time relation-aware representation generation. TiRGN [142] uses two encoders to mine the information at both local and global levels. HiSMatch [119] proposes different encoders to mine the semantic information of the historical query structures and candidate entities, respectively.

3.2.2   RNN-agnostic Model

RNN-agnostic models extend the original static KGR models by incorporating temporal information without RNN frameworks. According to how the time information guides the models, they can be roughly divided into two types, *i.e.,* time-vector guided models and time-operation guided models.

Time-Vector Guided Model. Time-vector guided models directly generate the additional temporal embedding **t** for temporal information and integrate them with the original fact embedding as additional information.

TComplEx and TNTComplEx [137] both come from ComplEx, where the fourth-order tensor space with additional consideration of time information is modeled by them. In the process of constructing subgraphs, the time embedding is used to calculate weighted probabilities in xERTE [116]. Then, T-GAP [162] encodes the query-specific structure patterns of Temporal KG and performs path-based reasoning based on it. CyGNet [118] attempts

to solve the entity prediction task by encoding the historical facts related to the subject entity in each query and the time-indexing vector is generated. ChronoR [124] builds on the basis of RotatE, which connects relation and time embeddings to obtain the overall rotation embedding applied to the final entity embedding. Furthermore, DBKGE [160] proposed an online inference algorithm that smoothed the representation vector of nodes over time. BoxTE [134] introduces a novel box representation method for temporal KGR based on the static KGR method BoxE. TuckERTNT [138] proposes a novel tensor decomposition model for Temporal KGs inspired by the Tucker decomposition of a 4-order tensor with the extra time embedding. TempoQR [136] generates question-specific time vectors and exploits these vectors to aggregate specific entities and their timestamps. TLT-KGE [130] captures semantic and time information as different axes of complex space. RotateQVS [144] aims to consider the time information changes with rotation operation in the latent space.

Time-Operation Guided Model. Time-operation guided models leverage some specific operations to combine the temporal information within the entity and relation embeddings instead of directly generating the temporal embedding **t**, such as encoding facts into designed time-specific hyper-planes and generating time-related rewards.

ChronoTranslate [159] learns a universal representation of entities and time-specific representations of the Temporal KGs, respectively. HyTE [157] represents each timestamp as a learnable hyper-plane in the embedding space, then projects entity and relation embeddings into the hyper-plane and utilizes the TransE scoring function on the projections. As a model for both graph learning and KGR, DyRep [153] captures the interleaved dynamics within history, which is further parameterized by a temporal-attentive representation network. Inspired by RotatE, to evaluate the given fact's semantic scores, TeRo [145] introduces a novel temporal guided rotation operation between head and tail entities. Diachronic embeddings [143] map entity and relation embeddings, paired with temporal information, into a KGR model space, thus defining a framework yielding specific models such as DE-TransE and DE-SimplE. Due to the uncertainty of temporal information in the graph evolution over time, ATiSE [141] maps the entity and relation embeddings of Temporal KGs into the Gaussian spaces according to the time stamps. TDGNN [139] introduces a novel temporal aggregator to combine the neighborhood features and the temporal information from edges to calculate the final representations. To mine the dynamic graph evolution of temporal KGs, DyERNIE [133] defines the velocity vector in the tangent space over time and encourages entity embeddings to evolve according to it. TPRec [158] is an interest recommendation method that presents an efficient time-aware interaction relation extraction component to construct a collaborative knowledge graph with time-aware interactions and also utilizes a time-aware path module for reasoning. TeLM [122] leverages a linear temporal regularizer and multi-vector encoders to realize the 4th-order tensor factorization for reasoning. RTFE [127] treats the Temporal KGs as a Markov chain, which transitions from the previous state to the next state, and then recursively tracks the state transition of Temporal KG by passing updated parameters/features between timestamps. Besides, TIE [120] combines the experience replay and time regularization into KGR to learn the time-aware incremental embedding. A length-aware CNN is leveraged in CEN [132] to handle historical facts via an easy-to-difficult

curriculum learning strategy over time. DKGE [140] introduces two different representations for each entity and each relationship (including temporal information). TLogic [128] performs the temporal random walks and extracts the temporal logical rules based on them, leading to better explainability. TKGC-AGP [126] leverages the approximations of multivariate Gaussian processes (MGPs) for fact encoding. Besides, FILT [125] makes use of the meta-learning framework for inferring the facts with unseen entities in temporal KGR. After that, rGalT [123] first designs the attention mechanism in both intra-graph and inter-graph levels to leverage the historical semantics. Similarly to it, DA-Net [117] also tries to learn attention weights on repetitive facts at different historical timestamps. MetaTKGR [121] dynamically adjusts the strategies of sampling and aggregating neighbors from recent facts for new entities through temporally supervised signals on future facts as instant feedback. Moreover, CENET [115] learns both the historical and non-historical dependency for inferring the most potential facts.

3.2.3 Discussion on Temporal KGR Models

Tab. 3 systematically presents a classification of the temporal KGR models, especially for the reasoning scenarios, *i.e.,* interpolation, and extrapolation scenarios. In summary, there are **29** interpolation reasoning models, including TTransE, ChronoTranslate, HyTE, TA-DISTMULT, DacKGR, DyERNIE, Diachronic embeddings, ATiSE, TeMP, ChronoR, TeLM, T-GAP, TIE, TPmod, DBKGE, Tpath, RTFE, DKGE, TuckERTNT, BoxTE, TempoQR, TLT-KGE, TRHyTE, ExKGR, RotateQVS, FILT, TKGC-AGP, TNT-ComplEX, TeRo. While there are **21** extrapolation reasoning models including CENET, CEN, Know-Evolve, xERTE, DyRep, TDGNN, EvolveGCN, RE-NET, HiSMatch, TPRec, CyGNet, CluSTeR, TimeTraveler, HIP, Tlogic, EvoKG, RE-GCN, DA-Net, MetaTKGR, rGalT, TiRGN.

According to it, we can further get the following observations, which may indicate the scope for different temporal KGR models and reveal the future trend in temporal KGR. (1) RNN-agnostic models, *i.e.,* time-vector guided models and time-operation guided models, generally treat the temporal information as additional attributes and integrate them into the previous static KGR models with different techniques. Such a manner is more flexible compared to RNN-based models. (2) Time-vector guided models encode the time information as an additional time vector **t**. Although these models are simple, their performances mostly rely on whether the time encoder and embedding fusion module are suitable. Unlike these models, time-operation guided models design specific time operations, which are task-specific. (3) RNN-based models can generally model the time information better than other models and can be more easily adopted to extrapolation scenarios. (4) The extrapolation reasoning is still at an early stage, occupying only around 30% of the temporal KGR models, which leaves space for further exploration.

3.3 Multi-Modal KGR Model

Existing multi-modal KGR models generally employ embedding-based reasoning models to infer the queried facts after fusing the multi-modal auxiliary features, *e.g.,* texts, images, etc. Because of the lack of fusion modules, directly applying static and temporal KGR models to multi-modal KGR tasks generally results in sub-optimal performance. As a research field at an early stage, the relevant works are not systematic enough [164], [165], [166]. To
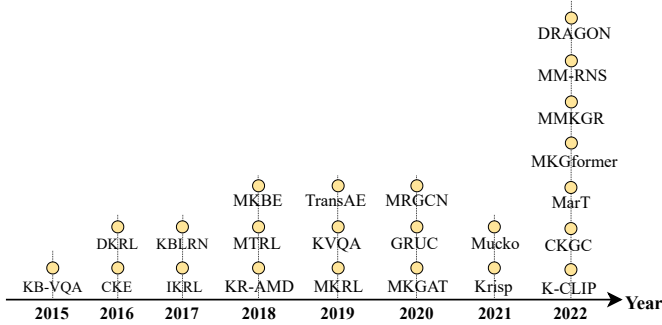
Figure 9: Timeline of the multi-modal KGR models.

this end, this section only comprehensively reviews **23** typical multi-modal KGR models along the timelines (See Fig. 9).

CKE [167] performs KG reasoning and collaborative filtering jointly, which enables it to generate representations and capture the implicit rules in KGs simultaneously. DKRL [168], [169] takes advantage of entity descriptions in KGs with language neural networks and gets more expressive semantics for reasoning. Inspired by it, IKRL [170] first designs an attention-based neural network to consider visual information in entity images. Such attention mechanism is also leveraged by TransAE [171]. KBLRN [172] first proposes an end-to-end reasoning framework, which combines neural network techniques with expert models for latent, relational, and numerical features. Besides, inspired by the translation-based static KGR models, MTRL [173] is a translation-based model with three energy functions corresponding to visual, linguistic, and structural information. Moreover, MKBE [174] and MRCGN [175] integrate different neural encoders and decoders with relational models for embedding learning and multi-modal data for reasoning. Then, KR-AMD [176] and MKRL [177] leverage textual data as part of auxiliary data to improve reasoning performance. MMKGR [178] first investigates the problem of how to effectively leverage multi-modal auxiliary features to conduct multi-hop reasoning in the KG area with a unified gate-attention network. MKGAT [179] better enhances recommendation systems with a multi-modal graph attention technique to conduct information propagation over multi-modal KGs. MarT [180] proposes a model-agnostic reasoning framework with a transformer for analogical reasoning. Besides, some hot techniques are integrated with multi-modal KGR models. For example, MM-RNS [181] and CKGC [182] leverage contrastive learning strategies, and a hybrid transformer with multi-level fusion is designed in MKGformer [183] designs. In the recent models, Multi-modal KGR techniques are also developed for the specific multi-modal downstream tasks to enhance the explainability of the models. KB-VQA [184] performs reasoning on the image and external knowledge, which provides an intuitive way to explain the generated answers. Similar to it, KVQA [185] integrates commonsense knowledge with images for reasoning. More recently, researchers have tended to first construct the scene KGs for a better understanding of the visual semantics, such as GRUC [186], Mucko [187], and Krisp [188]. Moreover, Knowledge-CLIP [189] leverages the CLIP [190] model for a better pre-trained model considering the semantic connections between multi-modal concepts. DRAGON [191] also provides a method for pre-training in a self-supervised manner for text and KG. As a research field at an early stage, there are still lots of space for deep exploration, which is further described in Sec. 5.4 as a challenge and opportunity.

3.4 Analysis on KGR models for Various Graph Types

Fig. 10 presents the statistic comparison of KGR models over static, temporal, and multi-modal KGs. According to it, we can find that static KGR models are over 50% proportions of the models. Compared to it, temporal KGR and multi-modal KGR models are relative in small proportions, revealing future trends.
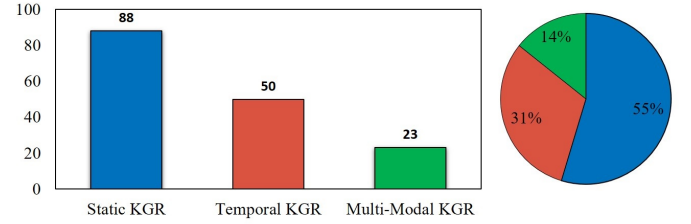

Figure 10: Statistic comparison of models over various KG types.

4 DATASETS

Our experiences show no comprehensive dataset collection for KGR tasks, especially for the temporal KGs and multi-modal KGs. To better convenience the community, we comprehensively summarize the datasets and systematically provide their statistics.

4.1 Static KGs

This section summarized the **38** transductive datasets for static KGR in the published papers. Moreover, **15** inductive datasets are proposed by [7] and [192] derived from the existing datasets, *i.e.,* WN18RR, FB15k-237, NELL-995. The statistic is presented in Tab. 4 and Tab. 5, and the descriptions are listed below.

- **ATMOIC** [193] is an KGs for everyday commonsense reasoning. It is composed of the reactions, effects, and intents of human behaviors and descriptions of each entity.
- **Countries** [194] consists of relations among countries based on public geographical data.
- **CoDEX** [195] is a set of COmpletion Datasets EXtracted from Wikidata and Wikipedia, which contains three sub-KGs according to the scope, *i.e.,* CoDEX-S, CoDEX-M, CoDEX-L. **CoDEX-S** can be used for model testing and debugging, as well as evaluation of methods that are less computationally efficient (*e.g.,* path-based and rule-based searching approaches). **CoDEX-M** [195] can be used for is all-purpose. Besides, **CoDEX-L** can be used for both general evaluations.
- **Conceptnet** [196] connects words and phrases with labeled edges to enhance AI application to better understand word meanings. Besides, **Conceptnet100K** [197] is the subset of it with 100k training triplets.
- **DBpedia** [226] consists of structured content from the information created in various Wikimedia projects. According to the entity set size, we can derive several subsets from it, *i.e.,* **DBpedia50** [198], **DBpedia500** [198] and **DB100K** [199].
- **FAMILY** [200] consists of relations among family members.
- **FreeBASE** [227] is a large knowledge base generated from multiple sources, such as Wikipedia, NNDB, Fashion Model Directory, etc. According to the entity set size, several subsets generated from it, including **FB13** [201], **FB122** [202], **FB15k** [203], **FB20k** [198], **FB24k** [204], **FB5M** [18], **FB15k-237** [205], **FB60k-NYT10** [206].

Table 4: Typical benchmark datasets for static transductive knowledge graph reasoning.

Dataset	# Entities	# Relations	# Train Triplets	# Val. Triplets	# Test Triplets
ATMOIC [193]	304,388	9	610,536	87,700	87,701
Countries [194]	271	2	1,110	24	24
CoDEX-S [195]	2,034	42	32,888	3,654	3656
CoDEX-M [195]	17,050	51	185,584	20620	20622
CoDEX-L [195]	77,951	69	551,193	30,622	30622
ConceptNet [196]	28,370,083	50	27,259,933	3,407,492	3,407,492
ConceptNet100K [197]	78,334	34	100,000	1,200	1,200
DBpedia50 [198]	49,900	654	32,388	399	10,969
DBpedia500 [198]	517,475	654	3,102,677	10,000	1,155,937
DB100K [199]	99,604	470	597,482	49,997	50,000
FAMILY [200]	3,007	12	23,483	2,038	2,835
FB13 [201]	75,043	13	316,232	11,816	47,464
FB122 [202]	9,738	122	91,638	9,595	11,243
FB15k [203]	14,951	1,345	483,142	50,000	59,071
FB20k [198]	19,923	1,345	472,860	48,991	90,149
FB24k [204]	23,634	673	402,493	-	21,067
FB5M [18]	5,385,322	1,192	19,193,556	50,000	59,071
FB15k-237 [205]	14,505	237	272,115	17,535	20,466
FB60k-NYT10 [206]	69,514	1,327	268,280	8,765	8,918
Hetionet [207]	45,158	24	1,800,157	225,020	225,020
Kinship [200]	104	25	8,544	1,068	1,074
Location [208]	445	5	384	65	65
Nation [209]	14	55	1,592	199	201
NELL23k [210]	22,925	200	25,445	4,961	4,952
NELL-995 [211]	75,492	200	126,176	5,000	5,000
OpenBioLink [212]	180,992	28	4,192,002	188,394	183,011
Sport [208]	1,039	4	1,349	358	358
Toy [213]	280	112	4,565	109	152
UMLS [214]	135	46	5,216	652	661
UMLS-PubMed [206]	59,226	443	2,030,841	8,756	8,689
WD-singer [210]	10,282	135	16,142	2,163	2,203
WN11 [201]	38,588	11	110,361	5,212	21,035
WN18 [215]	40,943	18	141,442	5,000	5,000
WN18RR [205]	40,559	11	86,835	2,924	2,824
wikidata5m [216]	4,594,485	822	20,614,279	5,163	5,163
YAGO3-10 [217]	123,143	37	1,079,040	4,978	4,982
YAGO37 [218]	123,189	37	420,623	50,000	50,000
M-/YAGO39k [219]	85,484	39	354,997	9,341	9,364

- **Hetionet** [207] is a knowledge graph derived from biomedical studies based on public resources. It describes relations among compounds, diseases, genes, anatomies, pathways, biological processes, molecular functions, cellular components, pharmacologic classes, side effects, and symptoms.
- **Kinship** [200] contains relations of the kinships in Alyawarra tribes [228].
- **Nation** [200] contains relations among nations [209].
- **NELL** [229] is the knowledge base built based on Never-Ending Language Learner, which attempts to learn to read the web over time. According to the different scopes, there are several subsets of it, *e.g.,* **Location** [208], **sports** [208], **NELL23k** [210], **NELL-995** [211].
- **OpenBioLink** [212] is a large-scale, high-quality, and highly challenging biomedical KG.
- **Toy** [213] is a small KG used for testing and debugging.
- **UMLS** [230] is the KG of the Unified Medical Language

System. By cooperating with the PubMed corpus, it is extended to **UMLS-PubMed** [206].
- **WordNet** [231] is a lexical database of semantic relations, such as synonyms, hyponyms, and meronyms, between words. According to the scope of the relations, there are some KGs derived from it, *e.g.,* **WN11** [201], **WN18** [215], **WN18RR** [205].
- **Wikidata** [232] provides a common source of data for Wikipedia. **WD-singer** [210] and **wikidata5m** [216] are derived from it according to different scopes.
- **YAGO** [233], [234], as a lightweight and extensible ontology, is built from **Wikidata** and unified with **WordNet**. According to the scopes of relations, **YAGO3-10** [217], **YAGO37** [218] and **YAGO39k** [219] are the subsets of it.

Table 5: Typical benchmark datasets for static inductive knowledge graph reasoning.

Dataset		# Entities	# Relations	# Train Triplets	# Val. Triplets	# Test Triplets
WN18RRv1 [7]	train-graph	2,746	9	5,410	626	638
	ind-test-graph	922	9	1,618	181	184
WN18RRv2 [7]	train-graph	6,954	10	15,262	1,837	1,868
	ind-test-graph	2,923	10	4,011	407	437
WN18RRv3 [7]	train-graph	12,078	11	25,901	3,097	3,152
	ind-test-graph	5,084	11	6,327	534	601
WN18RRv4 [7]	train-graph	3,861	9	7,940	934	968
	ind-test-graph	7,208	9	12,334	1,394	1,429
FB15k237v1 [7]	train-graph	2,000	183	4,245	485	492
	ind-test-graph	1,500	146	1,993	202	201
FB15k237v2 [7]	train-graph	3,000	203	9,739	1,166	1,180
	ind-test-graph	2,000	176	4,145	469	478
FB15k237v3 [7]	train-graph	4,000	218	17,986	2,194	2,214
	ind-test-graph	3,000	187	7,406	866	865
FB15k237v4 [7]	train-graph	5,000	222	27,203	3,352	3,361
	ind-test-graph	3,500	204	11,714	1,416	1,424
NELL995v1 [7]	train-graph	10,915	14	4,687	414	435
	ind-test-graph	225	14	833	97	96
NELL995v2 [7]	train-graph	2,564	88	8,219	922	968
	ind-test-graph	4,937	79	4,586	455	476
NELL995v3 [7]	train-graph	4647	142	16,393	1,851	1,873
	ind-test-graph	4,921	122	8,048	811	809
NELL995v4 [7]	train-graph	2,092	77	7,546	876	867
	ind-test-graph	3,294	61	7,073	716	731
WN-MBE [192]	train-graph	19,361	11	35,426	8,858	-
	ind-test-graph-1	3,723	11	5,678	-	1,352
	ind-test-graph-2	4,122	11	6,730	-	1,874
	ind-test-graph-3	4,300	11	7,545	-	2,054
	ind-test-graph-4	4467	11	8,623	-	2,493
	ind-test-graph-5	4,514	11	9,608	-	2,762
FB-MBE [192]	train-graph	7,203	237	125,769	31,442	-
	ind-test-graph-1	1,458	237	18,394	-	9,240
	ind-test-graph-2	1,461	237	19,120	-	9,669
	ind-test-graph-3	1,467	237	19,740	-	9,887
	ind-test-graph-4	1,467	237	22,455	-	11,127
	ind-test-graph-5	1,471	237	22,214	-	11,059
NELL-MBE [192]	train-graph	33,348	200	88,814	22,203	-
	ind-test-graph-1	34,488	3,200	34,496	-	3,853
	ind-test-graph-2	36031	3,200	35,411	-	31,059
	ind-test-graph-3	37,660	3,200	36,543	-	31,277
	ind-test-graph-4	39,056	3,200	37,667	-	31,427
	ind-test-graph-5	310,616	3,200	38,876	-	31,595

4.2 Temporal KGs

This section summarized the **18** benchmark datasets for temporal KGR collected from the published papers. The statistic of them are presented in Tab. 6, and the descriptions are listed below.

- **DBpedia-3SP** [220] is extracted subsets from **DBpedia** in three different timestamps.
- **GDELT** [130] is a dense KG derived from the Global Database of Events, Language, and Tone. **GDELT-m10** [221] and **GDELT-small** [120] is extracted from it.
- **IMDB** [235] is a KG consisting of the entities of movies, TV series, actors, and directors, which is also known as the Internet Movie Database. **IMDB-30SP** [130] and **IMDB-13-3SP** [130] are extracted from the dataset in different timestamps.
- **ICEWS** [236], short for Integrated Crisis Early Warning System, is a database that contains political events with specific timestamps. Some typical temporal KGs are created out of

it, *i.e.,* **ICEWS05-15** [222], **ICEWS11-14** [222], **ICEWS14** [116], **ICEWS14-Plus** [221], **ICEWS18** [116].
- **Wikidata** [232] for temporal KGR contains extra time information. **WIKI/Wikidata12k** [157], **Wikidata11k** [162] and **Wikidata-big** [136] are generated from it according to different periods.
- **YAGO** [233], [234] for temporal KGR contains extra time information. **YOGA11k/YOGA** [157], **YOGA15k** [222], **YOGA-3SP** [220] and **YOGA1830** [116] are generated from it according to different periods.

4.3 Multi-Modal KGs

This section summarized the **11** benchmark datasets for multi-modal KGR collected from the published papers. The statistic of them are presented in Tab. 7, and the descriptions are listed below.

- **FB-IMG-TXT** [173] is the KG combined with textual descriptions and images. The triple part is the subset of a

Table 6: Typical benchmark datasets for temporal knowledge graph reasoning.

Dataset	# Entities	# Relations	# Timestamps	# Train Triplets	# Val. Triplets	# Test Triplets
DBpedia-3SP [220]	66,967	968	3	103,211	3,000	-
GDELT [130]	7,691	240	8,925	1,033,270	238,765	305,241
GDELT-small [120]	500	20	366	2,735,685	341,961	341,961
GDELT-m10 [221]	50	20	30	221,132	27,608	27,926
IMDB-13-3SP [130]	3,244,455	14	3	7,913,773	10,000	-
IMDB-30SP [130]	243,148	14	30	621,096	3,000	3,000
ICEWS05-15 [222]	10,488	251	4,017	386,962	46,092	46275
ICEWS11-14 [222]	6,738	235	1,461	118,766	14,859	14,756
ICEWS14 [116]	7,128	230	365	63,685	13,823	13,222
ICEWS14-Plus [221]	7,128	230	365	72,826	8,941	8,963
ICEWS18 [116]	23,033	256	7,272	373,018	45,995	49,545
YOGA11k/YOGA [157]	10,623	10	189	161,540	19,523	20,026
YOGA-3SP [220]	27,009	37	3	124,757	3,000	3,000
YOGA15k [222]	15,403	34	198	110,441	13,815	13,800
YOGA1830 [116]	10,038	10	205	51,205	10,973	10,973
WIKI/Wikidata12k [157]	12,554	24	232	2,735,685	341,961	341,961
Wikidata11k [162]	11,134	95	328	242,844	28,748	14,283
Wikidata-big [136]	125,726	203	1,700	323,635	5,000	5,000

Table 7: Typical benchmark datasets for multi-modal knowledge graph reasoning.

Dataset	Modality	# Entities	# Relations	# Train Triplets	# Val. Triplets	# Test. Triplets
FB-IMG-TXT [173]	KG	11,757	1,231	285,850	34,863	29,580
	TXT	11,757				
	IMG	1,175,700				
FB15k-237-IMG [183]	KG	14,541	237	272,115	17,535	20,466
	IMG	145,410				
IMGpedia [223]	KG	14,765,300	442,959,000	3,119,207,705	-	-
	IMG	44,295,900				
MMKG-FB15k [224]	KG	14,951	1,345	592,213	-	-
	Numeric Literal	29,395		29,395	-	-
	IMG	13,444		13,444	-	-
MMKG-DB15k [224]	KG	14,777	279	99,028	-	-
	Numeric Literal	46,121		46,121	-	-
	IMG	12,841		12,841	-	-
MMKG-Yago15k [224]	KG	15,283	32	122,886	-	-
	Numeric Literal	48,405		48,405	-	-
	IMG	11,194		11,194	-	-
MKG-Wikipedia [181]	KG	15,000	169	34,196	4,274	4,276
	TXT	14,123				
	IMG	14,463				
MKG-YAGO [181]	KG	15,000	28	21,310	2,663	2,665
	TXT	12,305				
	IMG	14,244				
RichPedia [225]	KG	29,985	3	119,669,570	-	-
	IMG	2,914,770				
WN9-IMG-TXT [173]	KG	6,555	9	11,741	1,319	1,337
	TXT	6,555				
	IMG	63,225				
WN18-IMG [183]	KG	14,541	18	141,442	5,000	5,000
	IMG	145,410				

classical KG dataset FB15k [203], and the images are extracted from ImageNet [237]. Compared to it, **FB15K-237-IMG** [183] changes the scope of triplets to FB15k-237 [205].

- **IMGpedia** [223], [238], [239], [240] is the KG, which incorporates visual information of the images from the Wikimedia Commons dataset.
- **MKG** [181], [241] consists of two subsets, *i.e.,* **MKG-Wikipedia** and **MKG-YAGO**. They both contain visual entities generated by web search engines. But, their triplet parts are extracted from Wikipedia and YAGO respectively.
- **MMKG** [224] provides three subsets, including **MMKG-**

FB15k-IMG [224], **MMKG-DB15k** [224] and **Yago15k-IMG-TXT** [224], which integrates the specific KGs with numeric literals and images.

- **Richpedia** [225] is the KGs, composed of the triplets, textual descriptions, and images. The textual entities are derived from Wikidata, and the corresponding visual resources are extracted from the Web.
- **WN9-IMG-TXT** [170] is the KG combined with textual descriptions and images. The triple part is the subset of a classical KG dataset WN18 [215], and the images are extracted from ImageNet [237]. Compared to it, **WN18-IMG**

[183] changes the scope of triplets to the whole WN18 [215].

5 CHALLENGE AND OPPORTUNITY

According to previous analyses of the existing KGR models, we attempt to point out several promising directions for future works.

5.1 Out-of-distribution Reasoning

In the real-world scenario, new entities and relations are continuously emerging in the KGs, which are under-explored in the original KGs. Reasoning on the facts with these under-explored elements is called out-of-distribution reasoning, which raises higher requirements for the KGR model design. Some recent attempts provide potential solutions for inferring unseen entities, which are known as inductive reasoning models, such as [7], [70], [71], [73]. These models mine the logic rules underlying the graph structure without considering the specific meaning of entities, which achieve promising performances. As for the unseen relation inference, few-shot KGR models [72], [86], [121] tend to improve the generalization ability of models so that the trained model can scale well to the unseen relations with a small amount of facts. In other words, these few-shot KGR models can quickly learn new tasks according to the previously learned similar knowledge. Besides, BERTRL [84] tries to handle this case based on their textual semantics calculated by language models. While the performance of these models would drop drastically when language models are not finely trained. In conclusion, the KGR models for out-of-distribution reasoning tasks are still in an early stage, which is worth exploring in-depth in the future.

5.2 Large-scale Reasoning

The industrial KGs are generally large-scale, which requires more efficient KGR models. To this end, some existing works try to optimize the propagation procedures in a progressive manner [242]. For instance, NBF-net [81] integrates the bellman-ford algorithm to substitute the original DFS-based aggregation procedure in GNN-based KGR models. Moreover, A*Star [242] Net further optimizes the aggregation procedure with the greedy algorithm. Besides, the idea of graph clustering [243], [244], [245] is also used for it. For example, CURL [93] first separates the KGs into different clusters according to the entity semantics and then fine-grains the path-finding procedure into two-level, *i.e.,* the intra-cluster level and the inter-cluster level. It reduces the unnecessary searching for the whole graphs. Similarly, many works perform reasoning on sub-graphs instead of complete graphs, such as GraIL [7], CSR [86] etc. But most of them sacrifice the precision of inference, which may still be explored for more all-around models.

5.3 Multi-relational Reasoning

The situation that multi-relational facts exist between two entities is common in KGs as shown in Fig. 11 (a). However, they are more diverse in structure and more complex in semantics compared to uni-relational and bi-relational facts as shown in Figure 11 (b) and (c). Thus, the existing KGR models mainly focus on uni-relational and bi-relational facts and even usually treat multi-relational facts as uni-relational and bi-relational facts by omitting some of the facts. The KGR models in such a manner cannot accurately model real situations and lose lots of meaningful semantic information, leading to insufficient expressive ability. In the future, it is necessary to study how to leverage multi-relational facts to enhance reasoning ability.
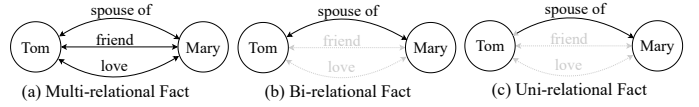

Figure 11: Comparison of multi-relational, bi-relational and uni-relational facts.

5.4 Multi-modal Reasoning

Knowledge reasoning based on the fusion of multi-source information can reduce the disconnectedness and sparsity of knowledge graphs by combining a text corpus or additional information in other modalities. Knowledge reasoning based on the fusion of data in multiple modalities can complement each other's advantages and improve reasoning performance. However, existing multi-modal KGR models are still at an early stage. They still tend to directly concat the embeddings in different modalities together for final score calculation. Such simple fusion modes have shown their promising performances while developing more fine-grained and scalable modes is still worthwhile. For instance, an adaptive fusion mode, which weighs the importance of different modalities, is worthwhile exploring.

5.5 Explainable Reasoning

Explainability is a common and important issue for deep learning models in various fields. Although KGR models generally are more explainable, it is still worthwhile exploring more in this topic, especially for embedding-based KGR models. Nowadays, more and more KGR models are developed based on neural networks, such as GNN. Most of them have the great expressive ability but suffer from explainability. Compared to them, rule-based and path-based KGR models are more explainable but computation-consuming and less expressive. To achieve a good trade-off between expressive ability and explainability, there exist some attempts to integrate the embedding-based models with rule-based and path-based models, such as ARGCN [192]. It builds the reward function based on the embeddings generated by the RGCN [57], which makes those path-based models more explainable. However, most of these attempts are still rough.

5.6 Knowledge Graph Reasoning Application

Although a large amount of KGR methods have been proposed in recent years, demonstrating the great potential of KGR in theoretical fields, the applications of KGR still need to be studied more. Nowadays, knowledge graphs are commonly used in many downstream applications, such as medicine, finance, plagiarism detection, etc. Medical knowledge reasoning models aim to assist doctors in diagnosing diseases from electronic medical records. For example, [246], [247] and [248] both perform reasoning on the KG constructed from the electronic medical database. The pre-trained language models, such as Bert, are leveraged to generate textual embedding of entities, which is proven effective in existing multi-modal KGR models. Besides, KGR models can also help with Anti-fraud detection, which is an important task in the finance field. For instance, [249] proposes a case-based reasoning method to assist people in verifying the information to discriminate against fraud in advance. Additionally, [250], [251] executes plagiarism detection by conducting the KGR approach in a continuous learning manner.

6 CONCLUSION

This survey systematically reviews 161 knowledge graph reasoning models for all types of knowledge graphs and comprehensively collects 67 typical datasets. Besides, the techniques and reasoning scenarios of the reviewed models are analyzed and discussed. Moreover, we point out some potential opportunities and list some promising future directions. The corresponding open-source repository is shared on GitHub: https://github.com/LIANGKE23/Awesome-Knowledge-Graph-Reasoning.

REFERENCES

[1] K. Liang, Y. Liu, S. Zhou, X. Liu, and W. Tu, "Relational symmetry based knowledge graph contrastive learning," 2022.

[2] E. H. Shortliffe, *Computer-Based Medical Consultation: MYCIN*. Elsevier North Holland, 1976.

[3] S. Ji, S. Pan, E. Cambria, P. Marttinen, and S. Y. Philip, "A survey on knowledge graphs: Representation, acquisition, and applications," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[4] C.-M. Wong, F. Feng, W. Zhang, C.-M. Vong, H. Chen, Y. Zhang, P. He, H. Chen, K. Zhao, and H. Chen, "Improving conversational recommender system by pretraining billion-scale knowledge graph," in *2021 IEEE 37th International Conference on Data Engineering (ICDE)*, 2021.

[5] G. Wang, W. Zhang, R. Wang, Y. Zhou, X. Chen, W. Zhang, H. Zhu, and H. Chen, "Label-free distant supervision for relation extraction via knowledge graph embedding," in *Proc. of EMNLP*, 2018.

[6] Y. Geng, J. Chen, Z. Chen, J. Z. Pan, Z. Ye, Z. Yuan, Y. Jia, and H. Chen, "Ontozsl: Ontology-enhanced zero-shot learning," in *Proc. of WWW*, 2021.

[7] K. Teru, E. Denis, and W. Hamilton, "Inductive relation prediction by subgraph reasoning," in *Proc. of ICML*, 2020.

[8] Z. Li, X. Jin, W. Li, S. Guan, J. Guo, H. Shen, Y. Wang, and X. Cheng, "Temporal knowledge graph reasoning based on evolutional representation learning," in *Proc. of SIGIR*, 2021.

[9] M. Liu and Y. Liu, "Inductive representation learning in temporal networks via mining neighborhood and community influences," in *Proc. of SIGIR*, 2021.

[10] M. Liu, J. Wu, and Y. Liu, "Embedding global and local influences for dynamic graphs," in *Proc. of CIKM*, 2022.

[11] P. Hitzler, K. Janowicz, G. Li, G. Qi, and Q. Ji, "Hybrid reasoning in knowledge graphs: Combing symbolic reasoning and statistical reasoning," *Semant. Web*, 2020.

[12] J. Zhang, B. Chen, L. Zhang, X. Ke, and H. Ding, "Neural, symbolic and neural-symbolic reasoning on knowledge graphs," *AI Open*, 2021.

[13] W. Zhang, J. Chen, J. Li, Z. Xu, J. Z. Pan, and H. Chen, "Knowledge graph reasoning with logics and embeddings: Survey and perspective," *arXiv preprint arXiv:2202.07412*, 2022.

[14] Y. Chen, H. Li, H. Li, W. Liu, Y. Wu, Q. Huang, and S. Wan, "An overview of knowledge graph reasoning: Key technologies and applications," *Journal of Sensor and Actuator Networks*, 2022.

[15] R. H. Richens, "Preprogramming for mechanical translation." *Mech. Transl. Comput. Linguistics*, 1956.

[16] X. Zhu, Z. Li, X. Wang, X. Jiang, P. Sun, X. Wang, Y. Xiao, and N. J. Yuan, "Multi-modal knowledge graph construction and application: A survey," *ArXiv*, 2022.

[17] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," *Proc. of NeurIPS*, 2013.

[18] Z. Wang, J. Zhang, J. Feng, and Z. Chen, "Knowledge graph embedding by translating on hyperplanes," in *Proc. of AAAI*, 2014.

[19] Y. Lin, Z. Liu, M. Sun, Y. Liu, and X. Zhu, "Learning entity and relation embeddings for knowledge graph completion," in *Proc. of AAAI*, 2015.

[20] G. Ji, S. He, L. Xu, K. Liu, and J. Zhao, "Knowledge graph embedding via dynamic mapping matrix," in *Proc. of ACL*, 2015.

[21] S. He, K. Liu, G. Ji, and J. Zhao, "Learning to represent knowledge graphs with gaussian embedding," in *Proceedings of the 24th ACM international on conference on information and knowledge management*, 2015.

[22] H. Xiao, M. Huang, Y. Hao, and X. Zhu, "Transg: A generative mixture model for knowledge graph embedding," *arXiv preprint arXiv:1509.05488*, 2015.

[23] G. Ji, K. Liu, S. He, and J. Zhao, "Knowledge graph completion with adaptive sparse transfer matrix," in *Proc. of AAAI*, 2016.

[24] T. Ebisu and R. Ichise, "Toruse: Knowledge graph embedding on a lie group," in *Proc. of AAAI*, 2018.

[25] I. Balazevic, C. Allen, and T. Hospedales, "Multi-relational poincaré graph embeddings," *Proc. of NeurIPS*, 2019.

[26] L. Ma, P. Sun, Z. Lin, and H. Wang, "Composing knowledge graph embeddings via word embeddings," *arXiv preprint arXiv:1909.03794*, 2019.

[27] Z. Sun, Z.-H. Deng, J.-Y. Nie, and J. Tang, "Rotate: Knowledge graph embedding by relational rotation in complex space," *arXiv preprint arXiv:1902.10197*, 2019.

[28] Z. Zhang, J. Cai, Y. Zhang, and J. Wang, "Learning hierarchy-aware knowledge graph embeddings for link prediction," in *Proc. of AAAI*, 2020.

[29] F. Zhang, X. Wang, Z. Li, and J. Li, "Transrhs: A representation learning method for knowledge graphs with relation hierarchical structure," in *Proc. of IJCAI*, 2021.

[30] L. Chao, J. He, T. Wang, and W. Chu, "Pairre: Knowledge graph embeddings via paired relation vectors," in *ACL*, 2021.

[31] R. Li, J. Zhao, C. Li, D. He, Y. Wang, Y. Liu, H. Sun, S. Wang, W. Deng, Y. Shen *et al.*, "House: Knowledge graph embedding with householder parameterization," *arXiv preprint arXiv:2202.07919*, 2022.

[32] L. Yu, Z. Luo, H. Liu, D. Lin, H. Li, and Y. Deng, "Triplere: Knowledge graph embeddings via tripled relation vectors," *arXiv preprint arXiv:2209.08271*, 2022.

[33] Y. Zhu, H. Liu, Z. Wu, Y. Song, and T. Zhang, "Representation learning with ordered relation paths for knowledge graph completion," 09 2019.

[34] M. Nickel, V. Tresp, and H.-P. Kriegel, "A three-way model for collective learning on multi-relational data," in *Icml*, 2011.

[35] B. Yang, W.-t. Yih, X. He, J. Gao, and L. Deng, "Embedding entities and relations for learning and inference in knowledge bases," *arXiv preprint arXiv:1412.6575*, 2014.

[36] T. Trouillon, J. Welbl, S. Riedel, É. Gaussier, and G. Bouchard, "Complex embeddings for simple link prediction," in *Proc. of ICML*, 2016.

[37] M. Nickel, L. Rosasco, and T. Poggio, "Holographic embeddings of knowledge graphs," in *Proc. of AAAI*, 2016.

[38] H. Liu, Y. Wu, and Y. Yang, "Analogical inference for multi-relational embeddings," in *Proc. of ICML*, 2017.

[39] S. M. Kazemi and D. Poole, "Simple embedding for link prediction in knowledge graphs," *Proc. of NeurIPS*, 2018.

[40] I. Balažević, C. Allen, and T. M. Hospedales, "Tucker: Tensor factorization for knowledge graph completion," *arXiv preprint arXiv:1901.09590*, 2019.

[41] W. Zhang, B. Paudel, W. Zhang, A. Bernstein, and H. Chen, "Interaction embeddings for prediction and explanation in knowledge graphs," in *Proc. of WSDM*, 2019.

[42] S. Zhang, Y. Tay, L. Yao, and Q. Liu, "Quaternion knowledge graph embeddings," *Proc. of NeurIPS*, 2019.

[43] Z. Cao, Q. Xu, Z. Yang, X. Cao, and Q. Huang, "Dual quaternion knowledge graph embeddings," in *Proc. of AAAI*, 2021.

[44] A. Bastos, K. Singh, A. Nadgeri, S. Shekarpour, I. O. Mulang, and J. Hoffart, "Hopfe: Knowledge graph representation learning using inverse hopf fibrations," in *Proc. of CIKM*, 2021.

[45] S. Amin, S. Varanasi, K. A. Dunfield, and G. Neumann, "Lowfer: Low-rank bilinear pooling for link prediction," in *Proc. of ICML*, 2020.

[46] D. Q. Nguyen, T. Vu, T. D. Nguyen, and D. Phung, "Quatre: Relation-aware quaternions for knowledge graph embeddings," in *Proc. of WWW*, 2022.

[47] A. Bordes, X. Glorot, J. Weston, and Y. Bengio, "A semantic matching energy function for learning with multi-relational data," *Machine Learning*, 2014.

[48] R. Socher, D. Chen, C. D. Manning, and A. Ng, "Reasoning with neural tensor networks for knowledge base completion," *Proc. of NeurIPS*, 2013.

[49] Q. Liu, H. Jiang, A. Evdokimov, Z.-H. Ling, X. Zhu, S. Wei, and Y. Hu, "Probabilistic reasoning via deep learning: Neural association models," *arXiv preprint arXiv:1603.07704*, 2016.

[50] B. Shi and T. Weninger, "Proje: Embedding projection for knowledge graph completion," in *Proc. of AAAI*, 2017.

[51] T. Dettmers, P. Minervini, P. Stenetorp, and S. Riedel, "Convolutional 2d knowledge graph embeddings," in *Proc. of AAAI*, 2018.

[52] D. Q. Nguyen, T. D. Nguyen, D. Q. Nguyen, and D. Phung, "A novel embedding model for knowledge base completion based on convolutional neural network," *arXiv preprint arXiv:1712.02121*, 2017.

[53] I. Balažević, C. Allen, and T. M. Hospedales, "Hypernetwork knowledge graph embeddings," in *Proc. of ICANN*, 2019.

[54] X. Jiang, Q. Wang, and B. Wang, "Adaptive convolution for multi-relational learning," in *Proc. of NAACL*, 2019.

[55] S. Vashishth, S. Sanyal, V. Nitin, N. Agrawal, and P. Talukdar, "Interacte: Improving convolution-based knowledge graph embeddings by increasing feature interactions," in *Proc. of AAAI*, 2020.

[56] C. Demir and A.-C. N. Ngomo, "Convolutional complex knowledge graph embeddings," in *European Semantic Web Conference*, 2021.

[57] M. Schlichtkrull, T. N. Kipf, P. Bloem, R. v. d. Berg, I. Titov, and M. Welling, "Modeling relational data with graph convolutional networks," in *European semantic web conference*, 2018.

[58] Z. Wang, Z. Ren, C. He, P. Zhang, and Y. Hu, "Robust embedding with multi-level structures for link prediction." in *Proc. of IJCAI*, 2019.

[59] D. Nathani, J. Chauhan, C. Sharma, and M. Kaul, "Learning attention-based embeddings for relation prediction in knowledge graphs," *arXiv preprint arXiv:1906.01195*, 2019.

[60] P. Wang, J. Han, C. Li, and R. Pan, "Logic attention based neighborhood aggregation for inductive knowledge graph embedding," in *Proc. of AAAI*, 2019.

[61] C. Shang, Y. Tang, J. Huang, J. Bi, X. He, and B. Zhou, "End-to-end structure-aware convolutional networks for knowledge base completion," in *Proc. of AAAI*, 2019.

[62] L. Cai, B. Yan, G. Mai, K. Janowicz, and R. Zhu, "Transgcn: Coupling transformation assumptions with graph convolutional networks for link prediction," in *Proceedings of the 10th International Conference on Knowledge Capture*, 2019.

[63] X. Xu, W. Feng, Y. Jiang, X. Xie, Z. Sun, and Z.-H. Deng, "Dynamically pruned message passing networks for large-scale knowledge graph reasoning," *arXiv preprint arXiv:1909.11334*, 2019.

[64] Z. Zhang, F. Zhuang, H. Zhu, Z. Shi, H. Xiong, and Q. He, "Relational graph neural network with hierarchical attention for knowledge graph completion," in *Proc. of AAAI*, 2020.

[65] D. Yu, Y. Yang, R. Zhang, and Y. Wu, "Knowledge embedding based graph convolutional network," in *Proc. of WWW*, 2021.

[66] S. Vashishth, S. Sanyal, V. Nitin, and P. Talukdar, "Composition-based multi-relational graph convolutional networks," *arXiv preprint arXiv:1911.03082*, 2019.

[67] J. Baek, D. B. Lee, and S. J. Hwang, "Learning to extrapolate knowledge: Transductive few-shot out-of-graph link prediction," *Proc. of NeurIPS*, 2020.

[68] Y. Zhang, W. Wang, W. Chen, J. Xu, A. Liu, and L. Zhao, "Meta-learning based hyper-relation feature modeling for out-of-knowledge-base embedding," in *Proc. of CIKM*, 2021.

[69] S. Liu, B. Grau, I. Horrocks, and E. Kostylev, "Indigo: Gnn-based inductive knowledge graph completion using pair-wise encoding," *Proc. of NeurIPS*, 2021.

[70] J. Chen, H. He, F. Wu, and J. Wang, "Topology-aware correlations between relations for inductive link prediction in knowledge graphs," in *Proc. of AAAI*, 2021.

[71] S. Mai, S. Zheng, Y. Yang, and H. Hu, "Communicative message passing for inductive relation reasoning." in *Proc. of AAAI*, 2021.

[72] S. Zheng, S. Mai, Y. Sun, H. Hu, and Y. Yang, "Subgraph-aware few-shot inductive link prediction via meta-learning," *IEEE Transactions on Knowledge and Data Engineering*, 2022.

[73] X. Xu, P. Zhang, Y. He, C. Chao, and C. Yan, "Subgraph neighboring relations infomax for inductive link prediction on knowledge graphs," *arXiv preprint arXiv:2208.00850*, 2022.

[74] Y. Pan, J. Liu, L. Zhang, X. Hu, T. Zhao, and Q. Lin, "Learning first-order rules with relational path contrast for inductive relation reasoning," *arXiv preprint arXiv:2110.08810*, 2021.

[75] H. Wang, H. Ren, and J. Leskovec, "Relational message passing for knowledge graph completion," in *Proc. of KDD*, 2021.

[76] Z. Hu, V. Gutiérrez-Basulto, Z. Xiang, X. Li, R. Li, and J. Z. Pan, "Type-aware embeddings for multi-hop reasoning over knowledge graphs," *arXiv preprint arXiv:2205.00782*, 2022.

[77] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Proc. of ICML*, 2020.

[78] Y. Zhu, Y. Xu, F. Yu, Q. Liu, S. Wu, and L. Wang, "Deep graph contrastive representation learning," *arXiv preprint arXiv:2006.04131*, 2020.

[79] Y. Liu, W. Tu, S. Zhou, X. Liu, L. Song, X. Yang, and E. Zhu, "Deep graph clustering via dual correlation reduction," in *Proc. of AAAI*, 2022.

[80] Y. Liu, X. Yang, S. Zhou, and X. Liu, "Simple contrastive graph clustering," *arXiv preprint arXiv:2205.07865*, 2022.

[81] Z. Zhu, Z. Zhang, L.-P. Xhonneux, and J. Tang, "Neural bellman-ford networks: A general graph neural network framework for link prediction," *Proc. of NeurIPS*, 2021.

[82] Y. Zhang and Q. Yao, "Knowledge graph reasoning with relational digraph," in *Proceedings of the ACM Web Conference 2022*, 2022.

[83] L. V. Harsha Vardhan, G. Jia, and S. Kok, "Probabilistic logic graph attention networks for reasoning," in *Proc. of WWW*, 2020.

[84] H. Zha, Z. Chen, and X. Yan, "Inductive relation prediction by bert," in *Proc. of AAAI*, 2022.

[85] Q. Lin, J. Liu, F. Xu, Y. Pan, Y. Zhu, L. Zhang, and T. Zhao, "Incorporating context graph with logical reasoning for inductive relation prediction," in *Proc. of SIGIR*, 2022.

[86] Q. Huang, H. Ren, and J. Leskovec, "Few-shot relational reasoning via connection subgraph pretraining," *arXiv preprint arXiv:2210.06722*, 2022.

[87] C. Fu, T. Chen, M. Qu, W. Jin, and X. Ren, "Collaborative policy learning for open knowledge graph reasoning," *arXiv preprint arXiv:1909.00230*, 2019.

[88] W. Zhang, B. Paudel, L. Wang, J. Chen, H. Zhu, W. Zhang, A. Bernstein, and H. Chen, "Iteratively learning embeddings and rules for knowledge graph reasoning," in *Proc. of WWW*, 2019.

[89] M. Qu and J. Tang, "Probabilistic logic neural networks for reasoning," *Proc. of NeurIPS*, 2019.

[90] A. Sadeghian, M. Armandpour, P. Ding, and D. Z. Wang, "Drum: End-to-end differentiable rule mining on knowledge graphs," *Proc. of NeurIPS*, 2019.

[91] P. G. Omran, K. Wang, and Z. Wang, "An embedding-based approach to rule learning in knowledge graphs," *IEEE Transactions on Knowledge and Data Engineering*, 2019.

[92] P.-W. Wang, D. Stepanova, C. Domokos, and J. Z. Kolter, "Differentiable learning of numerical rules in knowledge graphs," in *Proc. of ICLR*, 2019.

[93] D. Zhang, Z. Yuan, H. Liu, H. Xiong *et al.*, "Learning to walk with dual agents for knowledge graph reasoning," in *Proc. of AAAI*, 2022.

[94] H. Chen, Y. Li, S. Shi, S. Liu, H. Zhu, and Y. Zhang, "Graph collaborative reasoning," in *Proc. of WSDM*, 2022.

[95] Y. Shen, J. Chen, P.-S. Huang, Y. Guo, and J. Gao, "M-walk: Learning to walk in graph with monte carlo tree search," in *NIPS 2018*, 2018.

[96] X. V. Lin, R. Socher, and C. Xiong, "Multi-hop knowledge graph reasoning with reward shaping," *arXiv preprint arXiv:1808.10568*, 2018.

[97] W. Chen, W. Xiong, X. Yan, and W. Wang, "Variational knowledge graph reasoning," *arXiv preprint arXiv:1803.06581*, 2018.

[98] C. Meilicke, M. Fink, Y. Wang, D. Ruffinelli, R. Gemulla, and H. Stuckenschmidt, "Fine-grained evaluation of rule-and embedding-based systems for knowledge graph completion," in *Proc. of ISWC*, 2018.

[99] S. Guo, Q. Wang, L. Wang, B. Wang, and L. Guo, "Knowledge graph embedding with iterative guidance from soft rules," in *Proc. of AAAI*, 2018.

[100] R. Das, S. Dhuliawala, M. Zaheer, L. Vilnis, I. Durugkar, A. Krishnamurthy, A. Smola, and A. McCallum, "Go for a walk and arrive at the answer: Reasoning over paths in knowledge bases using reinforcement learning," *arXiv preprint arXiv:1711.05851*, 2017.

[101] W. Xiong, T. Hoang, and W. Y. Wang, "Deeppath: A reinforcement learning method for knowledge graph reasoning," *arXiv preprint arXiv:1707.06690*, 2017.

[102] T. Rocktäschel and S. Riedel, "End-to-end differentiable proving," *Proc. of NeurIPS*, 2017.

[103] F. Yang, Z. Yang, and W. W. Cohen, "Differentiable learning of logical rules for knowledge base reasoning," *Proc. of NeurIPS*, 2017.

[104] R. Das, A. Neelakantan, D. Belanger, and A. McCallum, "Chains of reasoning over entities, relations, and text using recurrent neural networks," *arXiv preprint arXiv:1607.01426*, 2016.

[105] S. Guo, Q. Wang, L. Wang, B. Wang, and L. Guo, "Jointly embedding knowledge graphs and logical rules," in *Proc. of EMNLP*, 2016.

[106] Y. Zhang, X. Chen, Y. Yang, A. Ramamurthy, B. Li, Y. Qi, and L. Song, "Efficient probabilistic logic reasoning with graph neural networks," *arXiv preprint arXiv:2001.11850*, 2020.

[107] A. Neelakantan, B. Roth, and A. McCallum, "Compositional vector space models for knowledge base completion," *arXiv preprint arXiv:1504.06662*, 2015.

[108] M. Gardner, P. Talukdar, J. Krishnamurthy, and T. Mitchell, "Incorporating vector space similarity in random walk inference over knowledge bases," in *Proc. of EMNLP*, 2014.

[109] L. A. Galárraga, C. Teflioudi, K. Hose, and F. Suchanek, "Amie: association rule mining under incomplete evidence in ontological knowledge bases," in *Proc. of WWW*, 2013.

[110] N. Lao and W. W. Cohen, "Relational retrieval using a combination of path-constrained random walks," *Machine learning*, 2010.

[111] G. F. Lawler and V. Limic, *Random walk: a modern introduction.* Cambridge University Press, 2010.

[112] K. Pearson, "The problem of the random walk," *Nature*, 1905.

[113] F. Spitzer, *Principles of random walk.* Springer Science & Business Media, 2001.

[114] J. Wang, J. Gao, Y. Yuan, and Q. Wang, "Crowd localization from gaussian mixture scoped knowledge and scoped teacher," 2022.

[115] Y. Xu, J. Ou, H. Xu, and L. Fu, "Temporal knowledge graph reasoning with historical contrastive learning," in *Proc. of AAAI*, 2022.

[116] Z. Han, P. Chen, Y. Ma, and V. Tresp, "Explainable subgraph reasoning for forecasting on temporal knowledge graphs," in *Proc. of ICLR*, 2020.

[117] K. Liu, F. Zhao, H. Chen, Y. Li, G. Xu, and H. Jin, "Da-net: Distributed attention network for temporal knowledge graph reasoning," in *Proc. of CIKM*, 2022.

[118] C. Zhu, M. Chen, C. Fan, G. Cheng, and Y. Zhang, "Learning from history: Modeling temporal knowledge graphs with sequential copy-generation networks," in *Proc. of AAAI*, 2021.

[119] Z. Li, Z. Hou, S. Guan, X. Jin, W. Peng, L. Bai, Y. Lyu, W. Li, J. Guo, and X. Cheng, "Hismatch: Historical structure matching based temporal knowledge graph reasoning," *arXiv preprint arXiv:2210.09708*, 2022.

[120] J. Wu, Y. Xu, Y. Zhang, C. Ma, M. Coates, and J. C. K. Cheung, "Tie: A framework for embedding-based incremental temporal knowledge graph completion," in *Proc. of SIGIR*, 2021.

[121] R. Wang, Z. Li, D. Sun, S. Liu, J. Li, B. Yin, and T. Abdelzaher, "Learning to sample and aggregate: Few-shot reasoning over temporal knowledge graphs," *arXiv preprint arXiv:2210.08654*, 2022.

[122] C. Xu, Y.-Y. Chen, M. Nayyeri, and J. Lehmann, "Temporal knowledge graph completion using a linear temporal regularizer and multivector embeddings," in *Proc. of NAACL*, 2021.

[123] Y. Gao, L. Feng, Z. Kan, Y. Han, L. Qiao, and D. Li, "Modeling precursors for temporal knowledge graph reasoning via auto-encoder structure," in *Proc. of IJCAI*, 2022.

[124] A. Sadeghian, M. Armandpour, A. Colas, and D. Z. Wang, "Chronor: rotation based temporal knowledge graph embedding," in *Proc. of AAAI*, 2021.

[125] Z. Ding, J. Wu, B. He, Y. Ma, Z. Han, and V. Tresp, "Few-shot inductive learning on temporal knowledge graphs using concept-aware information," *AKBC*, 2022.

[126] L. Zhang and D. Zhou, "Temporal knowledge graph completion with approximated gaussian process embedding," in *Proc. of COLING*, 2022.

[127] Y. Xu, E. Haihong, M. Song, W. Song, X. Lv, W. Haotian, and Y. Jinrui, "Rtfe: A recursive temporal fact embedding framework for temporal knowledge graph completion," in *Proc. of NAACL*, 2021.

[128] Y. Liu, Y. Ma, M. Hildebrandt, M. Joblin, and V. Tresp, "Tlogic: Temporal logical rules for explainable link forecasting on temporal knowledge graphs," in *Proc. of AAAI*, 2022.

[129] Y. He, P. Zhang, L. Liu, Q. Liang, W. Zhang, and C. Zhang, "Hip network: Historical information passing network for extrapolation reasoning on temporal knowledge graph." in *Proc. of IJCAI*, 2021.

[130] F. Zhang, Z. Zhang, X. Ao, F. Zhuang, Y. Xu, and Q. He, "Along the time: Timeline-traced embedding for temporal knowledge graph completion," in *Proc. of CIKM*, 2022.

[131] L. Bai, W. Yu, M. Chen, and X. Ma, "Multi-hop reasoning over paths in temporal knowledge graphs using reinforcement learning," *Applied Soft Computing*, 2021.

[132] Z. Li, S. Guan, X. Jin, W. Peng, Y. Lyu, Y. Zhu, L. Bai, W. Li, J. Guo, and X. Cheng, "Complex evolutional pattern learning for temporal knowledge graph reasoning," in *Proc. of ACL*, 2022.

[133] Z. Han, P. Chen, Y. Ma, and V. Tresp, "Dyernie: Dynamic evolution of riemannian manifold embeddings for temporal knowledge graph completion," in *Proc. of EMNLP*, 2020.

[134] J. Messner, R. Abboud, and I. I. Ceylan, "Temporal knowledge graph completion using box embeddings," in *Proc. of AAAI*, 2022.

[135] X. Lv, X. Han, L. Hou, J. Li, Z. Liu, W. Zhang, Y. Zhang, H. Kong, and S. Wu, "Dynamic anticipation and completion for multi-hop reasoning over sparse knowledge graph," in *Proc. of EMNLP*, 2020.

[136] C. Mavromatis, P. L. Subramanyam, V. N. Ioannidis, A. Adeshina, P. R. Howard, T. Grinberg, N. Hakim, and G. Karypis, "Tempoqr: temporal question reasoning over knowledge graphs," in *Proc. of AAAI*, 2022.

[137] T. Lacroix, G. Obozinski, and N. Usunier, "Tensor decompositions for temporal knowledge base completion," in *Proc. of ICLR*, 2019.

[138] P. Shao, D. Zhang, G. Yang, J. Tao, F. Che, and T. Liu, "Tucker decomposition-based temporal knowledge graph completion," *Knowledge-Based Systems*, 2022.

[139] L. Qu, H. Zhu, Q. Duan, and Y. Shi, "Continuous-time link prediction via temporal dependent graph neural network," in *Proc. of WWW*, 2020.

[140] T. Wu, A. Khan, M. Yong, G. Qi, and M. Wang, "Efficiently embedding dynamic knowledge graphs," *Knowledge-Based Systems*, 2022.

[141] C. Xu, M. Nayyeri, F. Alkhoury, H. Yazdi, and J. Lehmann, "Temporal knowledge graph completion based on time series gaussian embedding," in *Proc. of ISWC*, 2020.

[142] "Tirgn: Time-guided recurrent graph network with local-global historical patterns for temporal knowledge graph reasoning."

[143] R. Goel, S. M. Kazemi, M. Brubaker, and P. Poupart, "Diachronic embedding for temporal knowledge graph completion," in *Proc. of AAAI*, 2020.

[144] K. Chen, Y. Wang, Y. Li, and A. Li, "Rotateqvs: Representing temporal information as rotations in quaternion vector space for temporal knowledge graph completion," in *Proc. of ACL*, 2022.

[145] C. Xu, M. Nayyeri, F. Alkhoury, H. S. Yazdi, and J. Lehmann, "Tero: A time-aware knowledge graph embedding via temporal rotation," in *Proc. of COLING*, 2020.

[146] C. Yan, F. Zhao, and H. Jin, "Exkgr: Explainable multi-hop reasoning for evolving knowledge graph," in *Proc. of DASFAA*, 2022.

[147] A. Pareja, G. Domeniconi, J. Chen, T. Ma, T. Suzumura, H. Kanezashi, T. Kaler, T. Schardl, and C. Leiserson, "Evolvegcn: Evolving graph convolutional networks for dynamic graphs," in *Proc. of AAAI*, 2020.

[148] L. Yuan, Z. Li, J. Qu, T. Zhang, A. Liu, L. Zhao, and Z. Chen, "Trhyte: Temporal knowledge graph embedding based on temporal-relational hyperplanes," in *Proc. of DASFAA*, 2022.

[149] J. Wu, M. Cao, J. C. K. Cheung, and W. L. Hamilton, "Temp: Temporal message passing for temporal knowledge graph completion," in *Proc. of EMNLP*, 2020.

[150] N. Park, F. Liu, P. Mehta, D. Cristofor, C. Faloutsos, and Y. Dong, "Evokg: Jointly modeling event time and network structure for reasoning over temporal knowledge graphs," in *Proc. of WSDM*, 2022.

[151] W. Jin, M. Qu, X. Jin, and X. Ren, "Recurrent event network: Autoregressive structure inferenceover temporal knowledge graphs," in *Proc. of EMNLP*, 2020.

[152] L. Bai, X. Ma, M. Zhang, and W. Yu, "Tpmod: A tendency-guided prediction model for temporal knowledge graph completion," *ACM Transactions on Knowledge Discovery from Data*, 2021.

[153] R. Trivedi, M. Farajtabar, P. Biswal, and H. Zha, "Dyrep: Learning representations over dynamic graphs," in *Proc. of ICLR*, 2019.

[154] H. Sun, J. Zhong, Y. Ma, Z. Han, and K. He, "Timetraveler: Reinforcement learning for temporal knowledge graph forecasting," in *Proc. of EMNLP*, 2021.

[155] A. Garcia-Duran, S. Dumancic, and M. Niepert, "Learning sequence encoders for temporal knowledge graph completion," in *EMNLP*, 2018.

[156] Z. Li, X. Jin, S. Guan, W. Li, J. Guo, Y. Wang, and X. Cheng, "Search from history and reason for future: Two-stage reasoning on temporal knowledge graphs," in *ACL*, 2021.

[157] S. S. Dasgupta, S. N. Ray, and P. Talukdar, "Hyte: Hyperplane-based temporally aware knowledge graph embedding," in *Proc. of EMNLP*, 2018.

[158] Y. Zhao, X. Wang, J. Chen, Y. Wang, W. Tang, X. He, and H. Xie, "Time-aware path reasoning on knowledge graph for recommendation," *ACM Transactions on Information Systems (TOIS)*, 2021.

[159] A. Sadeghian, M. Rodriguez, D. Z. Wang, and A. Colas, "Temporal reasoning over event knowledge graphs," in *Workshop on Knowledge Base Construction, Reasoning and Mining*, 2016.

[160] S. Liao, S. Liang, Z. Meng, and Q. Zhang, "Learning dynamic embeddings for temporal knowledge graphs," in *Proc. of WSDM*, 2021.

[161] A. Garca-Duran, S. Dumancic, and M. Niepert, "Learning sequence encoders for temporal knowledge graph completion," in *EMNLP*, 2018.

[162] J. Jung, J. Jung, and U. Kang, "Learning to walk across time for interpretable temporal knowledge graph completion," in *Proc. of KDD*, 2021.

[163] R. Trivedi, H. Dai, Y. Wang, and L. Song, "Know-evolve: Deep temporal reasoning for dynamic knowledge graphs," in *Proc. of ICML*, 2017.

[164] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowledge-based systems*, 2013.

[165] X. Zhu, Z. Li, X. Wang, X. Jiang, P. Sun, X. Wang, Y. Xiao, and N. J. Yuan, "Multi-modal knowledge graph construction and application: A survey," *arXiv preprint arXiv:2202.05786*, 2022.

[166] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Advances in artificial intelligence*, 2009.

[167] F. Zhang, N. J. Yuan, D. Lian, X. Xie, and W.-Y. Ma, "Collaborative knowledge base embedding for recommender systems," in *Proc. of KDD*, 2016.

[168] R. Xie, Z. Liu, J. Jia, H. Luan, and M. Sun, "Representation learning of knowledge graphs with entity descriptions," in *Proc. of AAAI*, 2016.

[169] R. Xie, S. Heinrich, Z. Liu, C. Weber, Y. Yao, S. Wermter, and M. Sun, "Integrating image-based and knowledge-based representation learning," *IEEE Transactions on Cognitive and Developmental Systems*, 2019.

[170] R. Xie, Z. Liu, H. Luan, and M. Sun, "Image-embodied knowledge representation learning," in *Proc. of IJCAI*, 2017.

[171] Z. Wang, L. Li, Q. Li, and D. Zeng, "Multimodal data enhanced representation learning for knowledge graphs," in *Proc. of IJCNN*, 2019.

[172] A. Garcia-Duran and M. Niepert, "Kblrn: End-to-end learning of knowledge base representations with latent, relational, and numerical features," *arXiv e-prints*, 2017.

[173] H. Mousselly-Sergieh, T. Botschen, I. Gurevych, and S. Roth, "A multimodal translation-based approach for knowledge graph representation learning," in *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics*, 2018.

[174] P. Pezeshkpour, L. Chen, and S. Singh, "Embedding multimodal relational data for knowledge base completion," in *Proc. of EMNLP*, 2018.

[175] W. Wilcke, P. Bloem, V. de Boer, R. van t Veer, and F. van Harmelen, "End-to-end entity classification on multimodal knowledge graphs," *arXiv preprint arXiv:2003.12383*, 2020.

[176] Y. Zuo, Q. Fang, S. Qian, X. Zhang, and C. Xu, "Representation learning of knowledge graphs with entity attributes and multimedia descriptions," in *2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM)*, 2018.

[177] X. Tang, L. Chen, J. Cui, and B. Wei, "Knowledge representation learning with entity descriptions, hierarchical types, and textual relations," *Information Processing & Management*, 2019.

[178] S. Zheng, W. Wang, J. Qu, H. Yin, W. Chen, and L. Zhao, "Mmkgr: Multi-hop multi-modal knowledge graph reasoning," *arXiv preprint arXiv:2209.01416*, 2022.

[179] R. Sun, X. Cao, Y. Zhao, J. Wan, K. Zhou, F. Zhang, Z. Wang, and K. Zheng, "Multi-modal knowledge graphs for recommender systems," in *Proc. of CIKM*, 2020.

[180] N. Zhang, L. Li, X. Chen, X. Liang, S. Deng, and H. Chen, "Multimodal analogical reasoning over knowledge graphs," *arXiv preprint arXiv:2210.00312*, 2022.

[181] D. Xu, T. Xu, S. Wu, J. Zhou, and E. Chen, "Relation-enhanced negative sampling for multimodal knowledge graph completion," in *Proc. of ACM MM*, 2022.

[182] X. Cao, Y. Shi, J. Wang, H. Yu, X. Wang, and Z. Yan, "Cross-modal knowledge graph contrastive learning for machine learning method recommendation," in *Proc. of ACM MM*, 2022.

[183] X. Chen, N. Zhang, L. Li, S. Deng, C. Tan, C. Xu, F. Huang, L. Si, and H. Chen, "Hybrid transformer with multi-level fusion for multimodal knowledge graph completion," in *Proc. of SIGIR*, 2022.

[184] P. Wang, Q. Wu, C. Shen, A. v. d. Hengel, and A. Dick, "Explicit knowledge-based reasoning for visual question answering," *arXiv preprint arXiv:1511.02570*, 2015.

[185] S. Shah, A. Mishra, N. Yadati, and P. P. Talukdar, "Kvqa: Knowledge-aware visual question answering," in *Proc. of AAAI*, 2019.

[186] J. Yu, Z. Zhu, Y. Wang, W. Zhang, Y. Hu, and J. Tan, "Cross-modal knowledge reasoning for knowledge-based visual question answering," *Pattern Recognition*, 2020.

[187] Z. Zhu, J. Yu, Y. Wang, Y. Sun, Y. Hu, and Q. Wu, "Mucko: multi-layer cross-modal knowledge reasoning for fact-based visual question answering," in *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, 2021.

[188] K. Marino, X. Chen, D. Parikh, A. Gupta, and M. Rohrbach, "Krisp: Integrating implicit and symbolic knowledge for open-domain knowledge-based vqa," in *Proc. of CVPR*, 2021.

[189] X. Pan, T. Ye, D. Han, S. Song, and G. Huang, "Contrastive language-image pre-training with knowledge graphs," in *Proc. of NeurIPS*.

[190] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark *et al.*, "Learning transferable visual models from natural language supervision," in *Proc. of ICML*, 2021.

[191] M. Yasunaga, A. Bosselut, H. Ren, X. Zhang, C. D. Manning, P. Liang, and J. Leskovec, "Deep bidirectional language-knowledge graph pre-training," in *Proc. of NeurIPS*.

[192] Y. Cui, Y. Wang, Z. Sun, W. Liu, Y. Jiang, K. Han, and W. Hu, "Inductive knowledge graph reasoning for multi-batch emerging entities," in *Proc. of CIKM*, 2022.

[193] M. Sap, R. Le Bras, E. Allaway, C. Bhagavatula, N. Lourie, H. Rashkin, B. Roof, N. A. Smith, and Y. Choi, "Atomic: An atlas of machine commonsense for if-then reasoning," in *Proc. of AAAI*, 2019.

[194] G. Bouchard, S. Singh, and T. Trouillon, "On approximate reasoning capabilities of low-rank vector spaces," in *Proc. of AAAI*, 2015.

[195] T. Safavi and D. Koutra, "Codex: A comprehensive knowledge graph completion benchmark," in *Proc. of EMNLP*, 2020.

[196] R. Speer, J. Chin, and C. Havasi, "Conceptnet 5.5: An open multilingual graph of general knowledge," in *Proc. of AAAI*, 2017.

[197] J. Guo and S. Kok, "Bique: Biquaternionic embeddings of knowledge graphs," in *Proc. of EMNLP*, 2021.

[198] B. Shi and T. Weninger, "Open-world knowledge graph completion," in *Proc. of AAAI*, 2018.

[199] B. Ding, Q. Wang, B. Wang, and L. Guo, "Improving knowledge graph embedding using simple constraints," in *Proc. of ACL*, 2018.

[200] S. Kok and P. Domingos, "Statistical predicate invention," in *Proc. of ICML*, 2007.

[201] R. Socher, D. Chen, C. D. Manning, and A. Ng, "Reasoning with neural tensor networks for knowledge base completion," *Proc. of NeurIPS*, 2013.

[202] S. Guo, Q. Wang, L. Wang, B. Wang, and L. Guo, "Jointly embedding knowledge graphs and logical rules," in *Proc. of EMNLP*, 2016.

[203] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," *Proc. of NeurIPS*, 2013.

[204] Y. Lin, Z. Liu, and M. Sun, "Knowledge representation learning with entities, attributes and relations," in *Proc. of IJCAI*, 2016.

[205] K. Toutanova and D. Chen, "Observed versus latent features for knowledge base and text inference," in *Proceedings of the 3rd workshop on continuous vector space models and their compositionality*, 2015.

[206] C. Fu, T. Chen, M. Qu, W. Jin, and X. Ren, "Collaborative policy learning for open knowledge graph reasoning," in *Proc. of EMNLP*, 2019.

[207] D. S. Himmelstein, A. Lizee, C. Hessler, L. Brueggeman, S. L. Chen, D. Hadley, A. Green, P. Khankhanian, and S. E. Baranzini, "Systematic integration of biomedical knowledge prioritizes drugs for repurposing," *eLife*, 2017.

[208] S. Guo, Q. Wang, B. Wang, L. Wang, and L. Guo, "Semantically smooth knowledge graph embedding," in *Proc. of ACL*, 2015.

[209] R. J. Rummel, "Dimensionality of nations project," Tech. Rep., 1968.

[210] X. Lv, X. Han, L. Hou, J. Li, Z. Liu, W. Zhang, Y. Zhang, H. Kong, and S. Wu, "Dynamic anticipation and completion for multi-hop reasoning over sparse knowledge graph," in *Proc. of EMNLP*, 2020.

[211] W. Xiong, T. Hoang, and W. Y. Wang, "Deeppath: A reinforcement learning method for knowledge graph reasoning," in *EMNLP*, 2017.

[212] A. Breit, S. Ott, A. Agibetov, and M. Samwald, "OpenBioLink: a benchmarking framework for large-scale biomedical link prediction," *Bioinformatics*, 2020.

[213] S. Broscheit, D. Ruffinelli, A. Kochsiek, P. Betz, and R. Gemulla, "LibKGE - A knowledge graph embedding library for reproducible research," in *Proc. of EMNLP*, 2020.

[214] A. T. McCray, "An upper-level ontology for the biomedical domain," *Comparative and functional genomics*, 2003.

[215] K. Toutanova and D. Chen, "Observed versus latent features for knowledge base and text inference," in *Proceedings of the 3rd workshop on continuous vector space models and their compositionality*, 2015.

[216] X. Wang, T. Gao, Z. Zhu, Z. Zhang, Z. Liu, J. Li, and J. Tang, "Kepler: A unified model for knowledge embedding and pre-trained language representation," *Transactions of the Association for Computational Linguistics*, 2021.

[217] F. Mahdisoltani, J. Biega, and F. M. Suchanek, "A knowledge base from multilingual wikipedias–yago3," Tech. Rep., 2014.

[218] S. Guo, Q. Wang, L. Wang, B. Wang, and L. Guo, "Knowledge graph embedding with iterative guidance from soft rules," in *Proc. of AAAI*, 2018.

[219] X. Lv, L. Hou, J. Li, and Z. Liu, "Differentiating concepts and instances for knowledge graph embedding," in *Proc. of EMNLP*, 2018.

[220] T. Wu, A. Khan, M. Yong, G. Qi, and M. Wang, "Efficiently embedding dynamic knowledge graphs," *Knowledge-Based Systems*, 2022.

[221] Z. Han, G. Zhang, Y. Ma, and V. Tresp, "Time-dependent entity embedding is not all you need: A re-evaluation of temporal knowledge graph completion models under a unified framework," in *Proc. of EMNLP*, 2021.

[222] A. García-Durán, S. Dumancic, and M. Niepert, "Learning sequence encoders for temporal knowledge graph completion," in *EMNLP*, 2018.

[223] S. Ferrada, B. Bustos, and A. Hogan, "Imgpedia: a linked dataset with content-based analysis of wikimedia images," in *Proc. of ISWC*, 2017.

[224] Y. Liu, H. Li, A. Garcia-Duran, M. Niepert, D. Onoro-Rubio, and D. S. Rosenblum, "Mmkg: multi-modal knowledge graphs," in *European Semantic Web Conference*, 2019.

[225] M. Wang, G. Qi, H. Wang, and Q. Zheng, "Richpedia a comprehensive multi-modal knowledge graph," in *Joint International Semantic Technology Conference*, 2020.

[226] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. G. Ives, "Dbpedia: A nucleus for a web of open data," in *ISWC/ASWC*, 2007.

[227] K. D. Bollacker, C. Evans, P. K. Paritosh, T. Sturge, and J. Taylor, "Freebase: a collaboratively created graph database for structuring human knowledge," in *SIGMOD Conference*, 2008.

[228] W. W. Denham, "The detection of patterns in alyawara nonverbal behavior," Ph.D. dissertation, 1973.

[229] A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E. R. Hruschka, and T. M. Mitchell, "Toward an architecture for never-ending language learning," in *Proc. of AAAI*, 2010.

[230] O. Bodenreider, "The unified medical language system (umls): integrating biomedical terminology," *Nucleic acids research*, 2004.

[231] G. A. Miller, *WordNet: An electronic lexical database*. MIT press, 1998.

[232] D. Vrandečić and M. Krötzsch, "Wikidata: a free collaborative knowledgebase," *Communications of the ACM*, 2014.

[233] M. Fabian, K. Gjergji, W. Gerhard *et al.*, "Yago: A core of semantic knowledge unifying wordnet and wikipedia," in *Proc. of WWW*, 2007.

[234] F. M. Suchanek, G. Kasneci, and G. Weikum, "Yago: A large ontology from wikipedia and wordnet," *Journal of Web Semantics*, 2008.

[235] [Online]. Available: https://www.imdb.com/interfaces/

[236] E. Boschee, J. Lautenschlager, S. O'Brien, S. Shellman, J. Starz, and M. Ward, "ICEWS Coded Event Data," 2015. [Online]. Available: https://doi.org/10.7910/DVN/28075

[237] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Proc. of CVPR*, 2009.

[238] S. Ferrada, B. Bustos, and A. Hogan, "Imgpedia: Enriching the web of data with image content analysis." in *AMW*, 2016.

[239] B. Bustos and A. Hogan, "Imgpedia: a proposal to enrich dbpedia with image meta-data," in *Proc. 9 Alberto Mendelzon International Workshop on Foundations of Data Management (AMW)*, 2015.

[240] S. C. Ferrada Aliaga, "Imgpedia: A large-scale knowledge-base to perform visuo-semantic queries over wikimedia commons images," 2017.

[241] Z. Sun, Q. Zhang, W. Hu, C. Wang, M. Chen, F. Akrami, and C. Li, "A benchmarking study of embedding-based entity alignment for knowledge graphs," *Proceedings of the VLDB Endowment*, 2020.

[242] Y. Zhang, Z. Zhou, Q. Yao, X. Chu, and B. Han, "Learning adaptive propagation for knowledge graph reasoning," 2022.

[243] Y. Liu, J. Xia, S. Zhou, S. Wang, X. Guo, X. Yang, K. Liang, W. Tu, Z. S. Li, and X. Liu, "A survey of deep graph clustering: Taxonomy, challenge, and application," *arXiv preprint arXiv:2211.12875*, 2022.

[244] Y. Yang, Z. Guan, Z. Wang, W. Zhao, C. Xu, W. Lu, and J. Huang, "Self-supervised heterogeneous graph pre-training based on structural clustering," *arXiv preprint arXiv:2210.10462*, 2022.

[245] X. Yang, Y. Liu, S. Zhou, S. Wang, X. Liu, and E. Zhu, "Contrastive deep graph clustering with learnable augmentation," *arXiv preprint arXiv:2212.03559*, 2022.

[246] Y. Lan, S. He, K. Liu, X. Zeng, S. Liu, and J. Zhao, "Path-based knowledge reasoning with textual semantic information for medical knowledge graph completion," *BMC Medical Informatics and Decision Making*, 2021.

[247] K. Liang, S. Wu, and J. Gu, "Mka: A scalable medical knowledge-assisted mechanism for generative models on medical conversation tasks," in *Computational and Mathematical Methods in Medicine*, vol. vol. 2021, no. Article ID 5294627, 2021.

[248] M. Rotmensch, Y. Halpern, A. Tlimat, S. Horng, and D. Sontag, "Learning a health knowledge graph from electronic medical records," *Scientific reports*, 2017.

[249] S. Kapetanakis, G. Samakovitis, B. Gunasekara, and M. Petridis, "Monitoring financial transaction fraud with the use of case-based reasoning," 2012.

[250] M. Franco-Salvador, P. Gupta, P. Rosso, and R. E. Banchs, "Cross-language plagiarism detection over continuous-space- and knowledge graph-based representations of language," *Knowledge-Based Systems*, 2016.

[251] K. Liang, J. Tan, D. Zeng, Y. Huang, X. Huang, and G. Tan, "Abslearn: a gnn-based framework for aliasing and buffer-size information retrieval," *Pattern Analysis and Applications*, Feb 2023. [Online]. Available: https://doi.org/10.1007/s10044-023-01142-2