

# Universal Knowledge Graph Embeddings

N'Dah Jean Kouagou  
ndah.jean.kouagou@upb.de  
Department of Computer  
Science, Paderborn  
University  
Paderborn, NRW  
Germany

Caglar Demir  
caglar.demir@upb.de  
Department of Computer  
Science, Paderborn  
University  
Paderborn, NRW  
Germany

Hamada M. Zahera  
hamada.zahera@upb.de  
Department of Computer  
Science, Paderborn  
University  
Paderborn, NRW  
Germany

Adrian Wilke  
adrian.wilke@upb.de  
Department of Computer  
Science, Paderborn  
University  
Paderborn, NRW  
Germany

Stefan Heindorf  
heindorf@upb.de  
Department of Computer  
Science, Paderborn  
University  
Paderborn, NRW  
Germany

Jiayi Li  
jiayili@mail.uni-  
paderborn.de  
Department of Computer  
Science, Paderborn  
University  
Paderborn, NRW  
Germany

Axel-Cyrille Ngonga  
Ngomo  
axel.ngonga@upb.de  
Department of Computer  
Science, Paderborn  
University  
Paderborn, NRW  
Germany

## ABSTRACT

A variety of knowledge graph embedding approaches have been developed. Most of them obtain embeddings by learning the structure of the knowledge graph within a link prediction setting. As a result, the embeddings reflect only the structure of a single knowledge graph, and embeddings for different knowledge graphs are not aligned, e.g., they cannot be used to find similar entities across knowledge graphs via nearest neighbor search. However, knowledge graph embedding applications such as *entity disambiguation* require a more global representation, i.e., a representation that is valid across multiple sources. We propose to learn *universal knowledge graph embeddings* from large-scale interlinked knowledge sources. To this end, we fuse large knowledge graphs based on the owl:sameAs relation such that every entity is represented by a unique identity. We instantiate our idea by computing universal embeddings based on DBpedia and Wikidata yielding embeddings for about 180 million entities, 15 thousand relations, and 1.2 billion triples. We believe our computed embeddings will support the emerging field of graph foundation models. Moreover, we develop a convenient API to provide embeddings as a service. Experiments on link prediction suggest that *universal knowledge graph embeddings* encode better semantics compared to embeddings computed on a single knowledge graph. For reproducibility purposes, we provide our source code and datasets open access.<sup>1</sup>

<sup>1</sup>[https://github.com/dice-group/Universal\\_Embeddings](https://github.com/dice-group/Universal_Embeddings)



This work is licensed under a Creative Commons Attribution International 4.0 License.

## CCS CONCEPTS

• **Computing methodologies** → **Learning latent representations**; *Statistical relational learning*; *Neural networks*; • **Information systems** → **Entity relationship models**;

## KEYWORDS

Knowledge graph embedding, Universal KGE, Large KGs

### ACM Reference Format:

N'Dah Jean Kouagou, Caglar Demir, Hamada M. Zahera, Adrian Wilke, Stefan Heindorf, Jiayi Li, and Axel-Cyrille Ngonga Ngomo. 2024. Universal Knowledge Graph Embeddings. In *Companion Proceedings of the ACM Web Conference 2024 (WWW '24 Companion)*, May 13–17, 2024, Singapore, Singapore. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3589335.3651978>

## 1 INTRODUCTION

Knowledge graph embedding (KGE) models are used for a multitude of applications [4, 20], including link prediction, triple classification, entity classification, and entity resolution. Several approaches to compute KGEs have been proposed in the literature, e.g., TransE [2], RDF2Vec [12], and ComplEx [17].

While pretrained models for few knowledge graphs (KGs) are available, their embedding spaces are not aligned, i.e., same entities have different representations across different knowledge graphs. As a result, the usability of such embeddings is often limited to downstream tasks on the KG they were trained on [9]. However, a growing number of real-world applications of KG embeddings (e.g., graph foundation models [8, 10]) require entities to have a representation that integrates information from multiple sources.

The need for these unified representations for entities recently motivated several works [3, 14–16, 21, 24]. Some of these approaches are tailored towards multi-lingual KG embeddings, i.e., the task of computing aligned embeddings between multiple language versions of the same KG by relying on the available owl:sameAs links [3, 14, 21]. Other approaches employ a bootstrapping strategy based on the matching scores between entities during training [15]

or use additional information on entities such as attribute embeddings [16, 24]. Although current alignment approaches for KGs have shown promising results on benchmark datasets, they inherently suffer from scalability issues. This is corroborated by the lack of pretrained KG embeddings for large datasets such as Wikidata [18] and DBpedia [1] using the aforementioned approaches. Moreover, most entity alignment approaches can only handle two KGs at a time, and do not assign the same embedding vector to matching entities<sup>2</sup>.

In this paper, we merge a given set of KGs into a single KG to compute embeddings that capture comprehensive information about each entity. We call these embeddings *universal knowledge graph embeddings*. By assigning a unique ID to all matching entities, we not only reduce memory consumption and computation costs but also tackle KG incompleteness—a well-known issue in the research community [19, 25, 26]. For instance, Wikidata and DBpedia contain 360 and 138 triples about Iraq, respectively. Consequently, a traditional KGE model trained on either KG can only capture incomplete information about their shared entities. Our approach mitigates this limitation by integrating information from both KGs into the embeddings.

To quantify the quality of our embeddings, we apply our approach to 4 KGE models, i.e., DistMult[23], ComplEx[17], QMult [6], and ConEx [5], and evaluate their performance on link prediction. Overall, our results suggest that the benefits of using additional information derived from sameAs links become particularly noticeable in ConEx. We use the latter to compute and provide high-quality unified embeddings for the most populous KGs of the Linked Open Data (LOD) cloud<sup>3</sup>, i.e., DBpedia and Wikidata. The merged graph encompasses about 180 million entities, 15 thousand relations, and 1.2 billion triples. Moreover, we develop an API<sup>4</sup> with convenient methods to make the computed embeddings easily accessible.

## 2 RELATED WORK

A knowledge graph embedding (KGE) model denoted by  $\phi$  maps a knowledge graph (KG) into a continuous vector space—commonly by solving an optimization problem. This optimization aims to preserve the structural information of the input KG. For example, transitional distance models such as TransE [2] compute embeddings by modelling each triple as a translation between its head and tail entities. Specifically, TransE represents both entities and relations as vectors in the same semantic space and learns embeddings by minimizing the distance between  $h + r$  and  $t$  for every triple  $(h, r, t)$  in the considered KG. DistMult [23] adopts the scoring technique of TransE but uses multiplications to model entity and relation interactions. Other types of KG embedding models include ComplEx [17], ConEx [6] and RESCAL [11]. ComplEx and ConEx model entities and relations as complex vectors (i.e. with real and imaginary parts) to handle both symmetric and antisymmetric relations. As ComplEx cannot handle transitive relations (see Sun et al. 2019), ConEx further improves on ComplEx by applying a 2D convolution operation on complex-valued embeddings of head

<sup>2</sup>Two entities are matching if they correspond to the same real world entity

<sup>3</sup>LOD: <https://lod-cloud.net>

<sup>4</sup>API: <https://embeddings.cc/>

---

### Algorithm 1 Function FUSEKGs

---

**Input:**  $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_N$   
**Output:**  $\mathcal{G}^*$  # merged knowledge graphs

- 1:  $\mathcal{G}^* \leftarrow \mathcal{G}_1$  # Initialization
- 2: **for**  $i = 2, \dots, N$  **do**
- 3:   **for**  $(e_1, r, e_2) \in \mathcal{G}_i$  **do**
- 4:     **if**  $\exists e_1^* \in \mathcal{E}_{\mathcal{G}^*}$  such that  $e_1 : \text{sameAs} : e_1^*$  **then**
- 5:        $e_1 \leftarrow e_1^*$  # Rename entity  $e_1$
- 6:     **end if**
- 7:     **if**  $\exists e_2^* \in \mathcal{E}_{\mathcal{G}^*}$  such that  $e_2 : \text{sameAs} : e_2^*$  **then**
- 8:        $e_2 \leftarrow e_2^*$  # Rename entity  $e_2$
- 9:     **end if**
- 10:    Add  $(e_1, r, e_2)$  to  $\mathcal{G}^*$
- 11:   **end for**
- 12: **end for**
- 13: **return**  $\mathcal{G}^*$

---

entities and relations. By associating each relation with a matrix, RESCAL captures pairwise interactions between entities and is regarded as one of the most expressive models [20]. As KGs grow in size, computation-efficient algorithms are required to train KGs consisting of millions of entities and billions of triples. Zheng et al. 2020 developed DGL-KE, an open-source package that employs several optimization techniques to accelerate training on large KGs. For example, they partition a large KG to perform gradient updates on each partition and regularly fetch embeddings from other partitions which involves a significant communication overhead.

Although KG embeddings can benefit downstream tasks such as link prediction and KG completion, their successful application is often limited to the KGs they were trained on. As a result, applications to other tasks, such as *entity resolution on two or multiple KGs*, require aligned KG embeddings.

## 3 UNIVERSAL KNOWLEDGE GRAPH EMBEDDINGS

### 3.1 Preliminaries

A KG  $\mathcal{G}$  can be regarded as a set of triples  $\mathcal{G} = \{(e_1^{(i)}, r^{(i)}, e_2^{(i)})\}_{i=1}^n \subseteq \mathcal{E}_{\mathcal{G}} \times \mathcal{R}_{\mathcal{G}} \times \mathcal{E}_{\mathcal{G}}$ , where  $\mathcal{E}_{\mathcal{G}}$  and  $\mathcal{R}_{\mathcal{G}}$  represent its sets of entities and relations, respectively. When there is no ambiguity, we simply write  $\mathcal{E}$  and  $\mathcal{R}$ . Let  $\mathcal{G}_1, \dots, \mathcal{G}_N$  denote  $N$  KGs (in an arbitrary order), e.g., DBpedia, Wikidata, Freebase. Alignments between  $\mathcal{G}_i$  and  $\mathcal{G}_j$  are given by sameAs links. We use these links to fuse the given KGs as described in the next section.

### 3.2 Graph Fusion and Embedding Computation

In this work, we fuse all KGs  $\mathcal{G}_i$  for  $i = 1, \dots, N$  into a single KG  $\mathcal{G}^*$  where all aligned entities are represented by a unique ID. Algorithm 1 describes how the fusion is carried out. First, the algorithm chooses a reference KG<sup>5</sup> (in this work, we select  $\mathcal{G}_1$ ) as the initial set of triples for  $\mathcal{G}^*$  (line 1). Then, it iterates over the rest of the KGs (line 2) and adds their triples (line 10). In this process, entities that are already present in  $\mathcal{G}^*$  via sameAs links are renamed accordingly

<sup>5</sup>It does not matter which KG is chosen as we end up with the same number of triples in any case.

**Table 1: Statistics of the full datasets for *universal knowledge graph embeddings*. *Deg.* denotes the average degree of entities.**

Dataset	$ \mathcal{E} $	$ \mathcal{R} $	$ \mathcal{G} $	$ \text{sameAs} $	Deg.
DBpedia	91,684,304	13,783	616,564,603	33,860,047	13.45
Wikidata	94,468,182	1,436	667,666,110	33,860,047	14.14
MERGE	179,706,494	15,219	1,284,230,713	33,860,047	14.29

(lines 3–9). Once  $\mathcal{G}^*$  is constructed, a KGE model (e.g. ConEx) can be applied to learn universal embeddings.

### 3.3 Knowledge Graphs

We downloaded and preprocessed the September 2022 version of DBpedia [1] and the March 2022 version of Wikidata [18]. In this work, we only consider the English version of DBpedia, and its external links to Wikidata, i.e., `sameAs` links. The preprocessing step is concerned with the removal of non triplet-formatted files and literals.

*DBpedia ( $\mathcal{G}_1$ )*. DBpedia<sup>6</sup> is the most popular and prominent KG in the LOD. It is automatically created based on Wikipedia information, such as infobox tables, categorizations, and links to external websites. Since DBpedia serves as the hub for LOD, it contains many links to other LOD datasets such as Freebase, Caligraph, and Wikidata.

*Wikidata ( $\mathcal{G}_2$ )*. Wikidata<sup>7</sup> is a community-created knowledge base providing factual information to Wikipedia and other projects by the Wikimedia Foundation. As of April 2022, Wikidata contains over 97 million items and 1.37 billion statements. Each item page contains labels, short descriptions, aliases, statements, and site links. Each statement consists of a claim and an optional reference, and each claim consists of a property-value pair and optional qualifiers.

*Statistics of Knowledge Graphs*. Table 1 presents the statistics of DBpedia and Wikidata after our preprocessing step. MERGE is obtained by applying Algorithm 1 to {DBpedia, Wikidata}. In a normal scenario, the sum of the numbers of entities of DBpedia and Wikidata should be equal to that of MERGE and `sameAs` links. However, some entities in DBpedia were matched with multiple entities in Wikidata via `sameAs` links, and vice versa. This caused the equality not to hold as can be seen in the table.

## 4 EXPERIMENTS

We conduct our experiments to answer one fundamental question: “How do our universal knowledge embeddings compare to embeddings from traditional KGE approaches?” To this end, we set up a link prediction task where we compare two independently trained instances of the same embedding model (see the next sections for more details).

<sup>6</sup><https://databus.dbpedia.org/dbpedia/collections/dbpedia-snapshot-2022-09/>

<sup>7</sup><https://dumps.wikimedia.org/wikidatawiki/entities/>

**Table 2: Statistics of evaluation datasets. *Deg.* is the average degree of entities.**

Dataset	$ \mathcal{E} $	$ \mathcal{R} $	$ \mathcal{G} $	$ \text{sameAs} $	Deg.
DBpedia <sub>train</sub>	31,116	392	69,667	22,102	4.48
DBpedia <sub>test</sub>	15,602	279	15,374	10,471	1.97
Wikidata <sub>train</sub>	72,058	707	235,814	22,102	6.55
Wikidata <sub>test</sub>	41,137	465	53,761	10,471	2.61
MERGE <sub>train</sub>	81,836	1,099	305,481	22,102	7.47

### 4.1 Evaluation Setup

*Evaluation Datasets*. We conduct our experiments on subsets of DBpedia and Wikidata, due to the computational complexity of our evaluation metrics. Specifically, we randomly select 1% of entities in DBpedia that share `sameAs` links to Wikidata, then we obtain their 1-hop neighborhood together with the corresponding relation types. We then analogously compute the corresponding subset of Wikidata by using entities identified by the 1% initially selected in DBpedia. The samples we obtain are then randomly split into training and test datasets. Overall, we obtain five datasets for our experiments: DBpedia<sub>train</sub> and DBpedia<sub>test</sub> (split of the DBpedia sample), Wikidata<sub>train</sub> and Wikidata<sub>test</sub> (split of the Wikidata sample), and MERGE<sub>train</sub> (merge of DBpedia<sub>train</sub> and Wikidata<sub>train</sub> using Algorithm 1). For the sake of clarity, we use the notations DBpedia+ and Wikidata+ to refer to MERGE<sub>train</sub> depending on whether we are evaluating embedding models on DBpedia or Wikidata. Note that the splits are performed in a way that all entities and relation types in training datasets also appear in the test datasets. This allows us to not encounter out-of-vocabulary entities and relations at inference time. The sizes of the splits are specified in Table 2. The average degree abbreviated as *Deg.* represents the average number of edges connected to an entity.

*Metrics*. We use two standard metrics to evaluate KGEs: *hits@k* (*H@k*) and *mean reciprocal rank* (MRR). Formally, let  $\mathcal{G}$  be a knowledge graph, i.e., a set of triples. We denote by  $\text{rank}[e_1|r, e_2]$  the rank of the score of  $e_1$  given the relation  $r$  and the tail entity  $e_2$  among the set of all scores  $\{\text{score}(e|r, e_2) \text{ s. t. } e \in \mathcal{E}_{\mathcal{G}}\}$ . Similarly,  $\text{rank}[e_2|e_1, r]$  denotes the rank of the score of  $e_2$  given the head entity  $e_1$  and the relation  $r$  among  $\{\text{score}(e|e_1, r) \text{ s. t. } e \in \mathcal{E}_{\mathcal{G}}\}$ . We define the metrics *Hits@k* and MRR as

$$\text{Hits@k} = \frac{1}{2|\mathcal{G}|} \sum_{(e_1, r, e_2) \in \mathcal{G}} \mathbb{1}(\text{rank}[e_1|r, e_2] \leq k) + \mathbb{1}(\text{rank}[e_2|e_1, r] \leq k), \quad (1)$$

$$\text{MRR} = \frac{1}{2|\mathcal{G}|} \sum_{(e_1, r, e_2) \in \mathcal{G}} \frac{1}{\text{rank}[e_1|r, e_2]} + \frac{1}{\text{rank}[e_2|e_1, r]}. \quad (2)$$

*Hardware*. The entire DBpedia and Wikidata datasets for which we provide embeddings as a service were processed on a virtual machine (VM) with 128 CPUs (AMD EPYC 7742 64-Core Processor) and 1TB RAM. The computation of universal knowledge graph embeddings was carried out using the DICE embedding framework [7]

**Table 3: Link prediction results in terms of mean reciprocal rank (MRR) and hits@k (H@k). We compare the performance of each embedding model on the two types of datasets: single KG (DBpedia, Wikidata) and enriched KG (DBpedia+, Wikidata+). Hence, the bold values correspond to the best performance achieved row-wise. All models use 32 embedding dimensions and have approximately the same number of parameters.**

	DBpedia				DBpedia+			
	MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10
ConEx-train	0.371	0.298	0.396	0.502	<b>0.898</b>	<b>0.863</b>	<b>0.924</b>	<b>0.962</b>
ConEx-test	0.181	0.134	0.191	0.269	<b>0.349</b>	<b>0.280</b>	<b>0.378</b>	<b>0.478</b>
ComplEx-train	0.181	0.122	0.205	0.292	<b>0.278</b>	<b>0.206</b>	<b>0.306</b>	<b>0.408</b>
ComplEx-test	0.136	0.103	0.143	0.194	<b>0.163</b>	<b>0.123</b>	<b>0.170</b>	<b>0.237</b>
QMult-train	<b>0.230</b>	<b>0.168</b>	<b>0.250</b>	<b>0.348</b>	0.129	0.099	0.135	0.183
QMult-test	<b>0.148</b>	<b>0.112</b>	<b>0.155</b>	<b>0.213</b>	0.064	0.041	0.070	0.113
DistMult-train	<b>0.145</b>	<b>0.100</b>	<b>0.152</b>	<b>0.243</b>	0.125	0.094	0.133	0.183
DistMult-test	<b>0.118</b>	<b>0.084</b>	<b>0.125</b>	<b>0.177</b>	0.057	0.032	0.060	0.109
	Wikidata				Wikidata+			
	MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10
ConEx-train	0.795	0.730	0.842	0.911	<b>0.855</b>	<b>0.806</b>	<b>0.889</b>	<b>0.941</b>
ConEx-test	0.280	0.227	0.302	0.376	<b>0.299</b>	<b>0.247</b>	<b>0.324</b>	<b>0.392</b>
ComplEx-train	0.232	0.188	0.244	0.303	<b>0.278</b>	<b>0.206</b>	<b>0.306</b>	<b>0.408</b>
ComplEx-test	0.126	0.088	0.140	0.197	<b>0.166</b>	<b>0.123</b>	<b>0.181</b>	<b>0.247</b>
QMult-train	<b>0.138</b>	<b>0.105</b>	<b>0.147</b>	<b>0.199</b>	0.129	0.099	0.135	0.183
QMult-test	0.077	0.053	0.085	0.122	<b>0.092</b>	<b>0.069</b>	<b>0.100</b>	<b>0.135</b>
DistMult-train	<b>0.190</b>	<b>0.145</b>	<b>0.207</b>	<b>0.263</b>	0.125	0.094	0.133	0.183
DistMult-test	<b>0.107</b>	<b>0.075</b>	<b>0.118</b>	<b>0.165</b>	0.086	0.060	0.096	0.136

on a 1TB RAM VM with 2 NVIDIA RTX A5000 GPUs of 24GB memory each.

## 4.2 Results and Discussion

In Table 3, we present the results of our experiments comparing the performance of embedding models trained on a single KG against those trained on merged KGs (by leveraging sameAs links as described in Algorithm 1). Four embedding models are considered: ConEx [6], ComplEx [17], QMult [5] and DistMult [23]. From the table, we can observe that ConEx achieves the highest performance w.r.t. all metrics on all datasets. Moreover, its performance on the merged KGs (DBpedia+ and Wikidata+) is notably higher as compared to that on DBpedia and Wikidata. The ComplEx model also performs better on DBpedia+ and Wikidata+ than on DBpedia and Wikidata, respectively. On the other side, DistMult and QMult perform poorly on both the training and test datasets.

One would expect that additional information added about entities in DBpedia+ and Wikidata+ improves the performance of embedding models on downstream tasks such as link prediction. Although this is clearly the case for ConEx and ComplEx (with up to 2× improvement for ConEx), we observed the opposite on DistMult and QMult. Interestingly, the poor-performing models correspond to the extreme cases of model complexity, i.e., DistMult

is the simplest and QMult is the most expressive among the four models we considered. This suggests that with 32 embedding dimensions, DistMult cannot learn meaningful representations for entities and relations in our evaluation data due to its simplicity. Likewise, QMult fails to find optimal representations of entities and relations because it cannot encode its inherent high degree of freedom in 32 dimensions. The ConEx architecture appears to balance well between expressiveness and the chosen number of embedding dimensions. In fact, our preliminary experiments with 300 embedding dimensions ranked DistMult the top best-performing model ahead of ConEx and ComplEx at the cost of longer training times and memory consumption. In view of this observation, we use ConEx with 32 embedding dimensions to compute our universal embeddings for large KGs and provide them on a platform (see next section). The answer to the fundamental question behind our work is hence that *we can learn rich embeddings on a KG that integrates information (about entities) from different external sources, in particular other KGs*. A precondition to achieve this goal is intrinsic to common challenges in representation learning: find fitting hyper-parameters.

## 5 IMPLEMENTATION OF A SERVICE PLATFORM

We offer computed data as an open service, following the FAIR principles [22], so that the universal embeddings are available to a broad audience. The platform<sup>8</sup> consists of a RESTful, TLS-secured API along with a website and an interactive documentation. The API contains a hidden webservice for developers to maintain data and a public webservice with eight methods providing RDF entity identifiers and related embeddings.

Regarding the FAIR principles, the data is *findable* as existing entity identifiers are used and can be accessed and explored with an autocomplete feature on the website. The big data amount is *accessible* as data subsets can be retrieved using API methods such as *random* or *autocomplete*, which explore the full data. The accessibility is additionally enhanced by meta queries like the size of the offered datasets. *Interoperability* is given by reusing existing RDF namespaces and identifiers. In addition, the API is versioned and uses the lightweight JSON format in Python and JavaScript.

## 6 CONCLUSION AND OUTLOOK

In this paper, we discuss the challenges related to computing embeddings for entities shared across multiple knowledge graphs. In particular, we note the lack of such embeddings for large knowledge graphs and propose a simple but effective approach to compute embeddings for shared entities. Given a countable set of knowledge graphs, our approach iterates over all triples and assigns a unique ID to all matching entities (i.e. shared entities). An embedding model is then applied to learn embeddings on the resulting graph—our embeddings are called universal knowledge graph embeddings. We use our approach to compute embeddings for recent versions of DBpedia and Wikidata, and provide them as an open service via a convenient API. Experiments on link prediction suggest that our universal embeddings are better than those computed on separate knowledge graphs. Regarding the API, we currently

<sup>8</sup><https://embeddings.cc/>

provide embeddings via autocomplete search and random entity selection. In future releases, we will integrate an approximation of embedding-level nearest neighbour search to support real-time queries of similar entities over the complete data. We will also collect more large-scale knowledge graphs from the Linked Open Data Cloud to update our universal embeddings.

## ACKNOWLEDGEMENTS

This work has received funding from the European Union’s Horizon 2020 research and innovation programme within the project KnowGraphs under the Marie Skłodowska-Curie grant No 860801, the European Union’s Horizon Europe research and innovation programme within the project ENEXA under the grant No 101070305, and the European Union’s Horizon Europe research and innovation programme within the project LEMUR under the Marie Skłodowska-Curie grant agreement No 101073307. This work has also been supported by the Ministry of Culture and Science of North Rhine-Westphalia (MKW NRW) within the project SAIL under the grant No NW21-059D and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): TRR 318/1 2021 – 438445824.

## REFERENCES

- [1] Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary G. Ives. 2007. DBpedia: A Nucleus for a Web of Open Data. In *ISWC/ASWC*, Vol. 4825. Springer, 722–735.
- [2] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Durán, Jason Weston, and Oksana Yakhnenko. 2013. Translating Embeddings for Modeling Multi-relational Data. In *NIPS*. 2787–2795.
- [3] Muhao Chen, Yingtao Tian, Mohan Yang, and Carlo Zaniolo. 2017. Multilingual Knowledge Graph Embeddings for Cross-lingual Knowledge Alignment. In *IJCAI*. 1511–1517.
- [4] Yuanfei Dai, Shiping Wang, Neal N Xiong, and Wenzhong Guo. 2020. A survey on knowledge graph embedding: Approaches, applications and benchmarks. *Electronics* 9, 5 (2020), 750.
- [5] Caglar Demir, Diego Moussallem, Stefan Heindorf, and Axel-Cyrille Ngonga Ngomo. 2021. Convolutional Hypercomplex Embeddings for Link Prediction. In *ACML (Proceedings of Machine Learning Research, Vol. 157)*. PMLR, 656–671.
- [6] Caglar Demir and Axel-Cyrille Ngonga Ngomo. 2021. Convolutional Complex Knowledge Graph Embeddings. In *ESWC*, Vol. 12731. Springer, 409–424.
- [7] Caglar Demir and Axel-Cyrille Ngonga Ngomo. 2022. Hardware-agnostic computation for large-scale knowledge graph embeddings. *Softw. Impacts* 13 (2022), 100377.
- [8] Mikhail Galkin, Xinyu Yuan, Hesham Mostafa, Jian Tang, and Zhaocheng Zhu. 2023. Towards Foundation Models for Knowledge Graph Reasoning. *CoRR* abs/2310.04562 (2023).
- [9] Nitisha Jain, Jan-Christoph Kalo, Wolf-Tilo Balke, and Ralf Krestel. 2021. Do Embeddings Actually Capture Knowledge Graph Semantics?. In *ESWC*, Vol. 12731. Springer, 143–159.
- [10] Jiawei Liu, Cheng Yang, Zhiyuan Lu, Junze Chen, Yibo Li, Mengmei Zhang, Ting Bai, Yuan Fang, Lichao Sun, Philip S. Yu, and Chuan Shi. 2023. Towards Graph Foundation Models: A Survey and Beyond. *CoRR* abs/2310.11829 (2023).
- [11] Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. 2011. A Three-Way Model for Collective Learning on Multi-Relational Data. In *ICML*. Omnipress, 809–816.
- [12] Petar Ristoski and Heiko Paulheim. 2016. RDF2Vec: RDF Graph Embeddings for Data Mining. In *ISWC*, Vol. 9981. 498–514.
- [13] Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. 2019. RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space. In *ICLR (Poster)*. OpenReview.net.
- [14] Zequn Sun, Wei Hu, and Chengkai Li. 2017. Cross-Lingual Entity Alignment via Joint Attribute-Preserving Embedding. In *ISWC (1) (Lecture Notes in Computer Science, Vol. 10587)*. Springer, 628–644.
- [15] Zequn Sun, Wei Hu, Qingheng Zhang, and Yuzhong Qu. 2018. Bootstrapping Entity Alignment with Knowledge Graph Embedding. In *IJCAI*. ijcai.org, 4396–4402.
- [16] Bayu Distiawan Trisedya, Jianzhong Qi, and Rui Zhang. 2019. Entity Alignment between Knowledge Graphs Using Attribute Embeddings. In *AAAI*. AAAI Press, 297–304.
- [17] Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard. 2016. Complex Embeddings for Simple Link Prediction. In *ICML (JMLR Workshop and Conference Proceedings, Vol. 48)*. JMLR.org, 2071–2080.
- [18] Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Commun. ACM* 57, 10 (2014), 78–85.
- [19] Meng Wang, Haomin Shen, Sen Wang, Lina Yao, Yinlin Jiang, Guilin Qi, and Yang Chen. 2019. Learning to Hash for Efficient Search Over Incomplete Knowledge Graphs. In *ICDM*. IEEE, 1360–1365.
- [20] Quan Wang, Zhendong Mao, Bin Wang, and Li Guo. 2017. Knowledge Graph Embedding: A Survey of Approaches and Applications. *IEEE Trans. Knowl. Data Eng.* 29, 12 (2017), 2724–2743.
- [21] Zhichun Wang, Qingsong Lv, Xiaohan Lan, and Yu Zhang. 2018. Cross-lingual Knowledge Graph Alignment via Graph Convolutional Networks. In *EMNLP*. Association for Computational Linguistics, 349–357.
- [22] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3, 1 (2016), 1–9.
- [23] Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. 2015. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. In *ICLR (Poster)*.
- [24] Qingheng Zhang, Zequn Sun, Wei Hu, Muhao Chen, Lingbing Guo, and Yuzhong Qu. 2019. Multi-view Knowledge Graph Embedding for Entity Alignment. In *IJCAI*. 5429–5435.
- [25] Xiaoyu Zhang, Xin Xin, Dongdong Li, Wenxuan Liu, Pengjie Ren, Zhumin Chen, Jun Ma, and Zhaochun Ren. 2023. Variational Reasoning over Incomplete Knowledge Graphs for Conversational Recommendation. In *WSDM*. ACM, 231–239.
- [26] Fen Zhao, Yinguo Li, Jie Hou, and Ling Bai. 2022. Improving question answering over incomplete knowledge graphs with relation prediction. *Neural Comput. Appl.* 34, 8 (2022), 6331–6348.
- [27] Da Zheng, Xiang Song, Chao Ma, Zeyuan Tan, Zihao Ye, Jin Dong, Hao Xiong, Zheng Zhang, and George Karypis. 2020. DGL-KE: Training Knowledge Graph Embeddings at Scale. In *SIGIR*. ACM, 739–748.