Beyond Vector Proximity: Evaluating Semantic Cohesion in Knowledge Graph Embeddings

Verdiana Schena[1,2], Simona Colucci[1], Francesco M. Donini[3], Catia Pesquita[2] and Floriano Scioscia[1]

[1]Polytechnic University of Bari, Bari, Italy
[2]University of Lisbon, Lisbon, Portugal
[3]University of Tuscia, Viterbo, Italy

Abstract

Knowledge graph embedding models represent entities as vectors in a latent space, where geometric proximity is often assumed to reflect semantic similarity. However, this proximity is not directly interpretable, since the dimensions and geometric relations of the embedding space have no explicit correspondence with the symbolic features of the underlying knowledge graph. Assessing whether nearby embeddings are associated with identifiable semantic commonalities can therefore provide a basis for interpreting proximity-based outputs, such as entity neighborhoods and clusters. In this work, we investigate the relationship between geometric cohesion in embedding spaces and semantic similarity in knowledge graphs at two complementary levels. At the pairwise level, we assess whether the proximity between individual entity embeddings can be grounded in explicit semantic evidence by comparing cosine similarity with the Partitioned Information Content-Based Semantic Similarity measure, PICSS. At the groupwise level, we introduce LD-SimGIC, a semantic similarity measure for assessing the semantic cohesion of resource clusters in Linked Data, and correlate it with the geometric cohesion of the corresponding embedding clusters. This distinction allows us to examine both local semantic alignment between pairs of resources and broader semantic regularities emerging within groups. We evaluate several knowledge graph embedding models on two different datasets: DrugBank and SemOpenAlex. Our results reveal a clear discrepancy between the two settings: pairwise correlation is very weak for almost all models, whereas strong groupwise correlation emerges for a subset of models, most consistently CrossE, DistMA, NTN, PairRE, ProjE, SE, TransF and UM. Finally, we complement the quantitative analysis with natural-language, logic-based explanations of the features shared by nearby resources, in an effort toward user-centered interpretability.

Keywords

Knowledge Graph Embeddings, Semantic Similarity, Linked Data, Groupwise Semantic Similarity.

CEUR-WS.org/Vol-4318/paper3.pdf

1. Introduction

Knowledge graph embedding models (KGEMs) are commonly expected to capture semantic information [1], such that semantically similar entities are mapped to nearby points in the latent embedding space [2, 3, 4]. However, the notion of entity similarity lacks a universally accepted definition in knowledge graphs (KGs) [3], since semantic similarity can be characterized from different perspectives [5], including structural and contextual ones. This makes it non-trivial to assess whether geometric proximity in an embedding space actually reflects meaningful semantic similarity between entities. Moreover, geometric proximity is not directly interpretable, since neither the dimensions of the embeddings nor the geometric relations between entity representations have an explicit correspondence with the symbolic features of the underlying knowledge graph [2, 6]. Assessing whether nearby embeddings are associated with identifiable semantic commonalities can therefore provide a basis for interpreting proximity-based outputs, such as entity neighborhoods and clusters.

In this work, we investigate the relationship between geometric cohesion in embedding spaces and

XAI-KG@ISWC2026 - 2nd International Workshop on eXplainable AI and Knowledge Graphs, co-located with ISWC 2026, October 25-29, 2026, Bari, Italy

v.schena@phd.poliba.it (V. Schena); simona.colucci@poliba.it (S. Colucci); donini@unitus.it (F. M. Donini); clpesquita@ciencias.ulisboa.pt (C. Pesquita); floriano.scioscia@poliba.it (F. Scioscia)
0009-0009-1660-5462 (V. Schena); 0000-0002-8774-4816 (S. Colucci); 0000-0003-0284-9625 (F. M. Donini); 0000-0002-1847-9393 (C. Pesquita); 0000-0002-7859-9602 (F. Scioscia)

© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2026-11-07

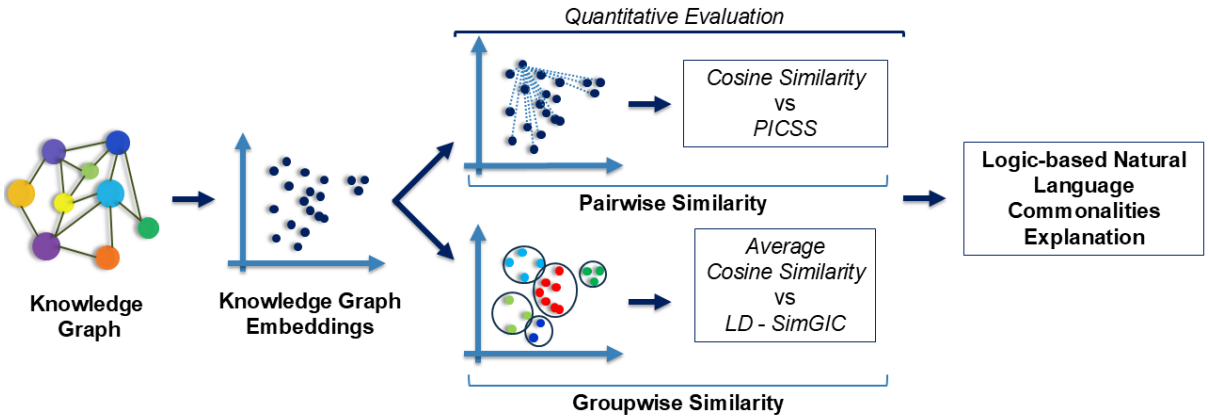


Figure 1: Overview of the proposed workflow, from entity embedding generation to pairwise and groupwise analysis, and finally to commonalities explanation.

semantic similarity derived from the corresponding knowledge graph resources at two complementary levels. At the pairwise level, we examine whether the proximity between two entity embeddings is aligned with explicit semantic evidence from the knowledge graph. This local analysis is particularly relevant when embeddings are used in downstream tasks that rely on individual entity neighborhoods, such as recommendation [7] and entity alignment [8]. We therefore compare geometric similarity between resource pairs with their semantic similarity computed from Linked Data (LD) features, i.e., incoming and outgoing RDF relations connecting a resource to other URI-identified resources. We then extend the analysis from pairs to groups of resources. In particular, we use clustering to study whether sets of entities that are geometrically coherent in the embedding space are also semantically coherent. This groupwise analysis considers a broader level of semantic organization and is relevant to the interpretation of embedding-induced clusters, which may be used in tasks such as exploratory analysis and clustering [9, 4, 10]. In both settings, the analysis correlates geometric cohesion with semantic similarity derived from LD features and weighted by their information content (IC), which quantifies their specificity in the KG, assigning higher values to less frequent and therefore more informative features [11]. The pairwise and groupwise analyses thus provide complementary views: the former evaluates local semantic alignment between individual resources, whereas the latter evaluates whether broader semantic regularities emerge within geometrically cohesive groups. Finally, we complement the quantitative analysis with natural language (NL), logic-based explanations that make explicit the features shared by nearby resources, in an effort toward user-centered interpretability. Figure 1 provides an overview of the proposed workflow.

The remainder of the paper is organized as follows. Section 2 reviews related work on semantic similarity measures and on the ability of knowledge graph embeddings to capture semantics, while positioning our contribution with respect to the existing literature. Section 3 presents the proposed methodology by describing: how geometric and semantic similarity are computed in both the pairwise and groupwise settings, how correlation analysis is performed, and how natural language, logic-based explanations are generated to make explicit what nearby resources in the latent space actually share. Section 4 describes the experimental setting and implementation details. Section 5 reports and discusses the results. Finally, Section 6 concludes the paper and outlines directions for future work.

2. Related Work

2.1. Semantic Similarity in Ontologies and Linked Data

Semantic similarity is generally grounded in two complementary intuitions: it increases with the amount of information shared by entities and decreases with the amount of information that distinguishes them [11, 12]. Accordingly, measures based only on commonalities implicitly assume an initial similarity

of zero and increase it with shared features, whereas difference-aware measures also decrease it in the presence of distinctive features [5]. Over the years, many semantic similarity measures have been proposed, mostly in the ontology domain. These approaches can be broadly categorized into edge-based, node-based, and hybrid methods, with node-based approaches further divided into feature-based and IC-based measures [13]. An additional classification distinguishes between pairwise and groupwise similarity measures [14]. While pairwise approaches compare two concepts at a time, groupwise measures, such as SimGIC [15], enable the comparison of sets of concepts. More recently, the focus has shifted from ontology-based settings to Resource Description Framework (RDF) graphs and LD, where resources are not organized only through hierarchical *is-a* relations, but are connected by heterogeneous links expressing a wide range of relation types. In this context, many ontology-based similarity measures, such as edge-counting or most informative common ancestor-based approaches, cannot be directly applied [16]. To address these limitations, Meymandpour and Davis [16] proposed the Partitioned Information Content-based Semantic Similarity measure (PICSS), a pairwise semantic similarity measure specifically designed for LD. PICSS represents each resource using the features extracted from its incoming and outgoing RDF links and computes similarity by considering not only the features shared by two resources but also their information content. While PICSS provides an information content-based formulation for comparing pairs of RDF resources, it does not directly support the assessment of semantic cohesion within groups of resources. Conversely, SimGIC enables groupwise similarity assessment, but it is defined for ontology-based annotations and cannot be directly applied to LD, where resources are described through heterogeneous RDF relations. To address these complementary limitations, we introduce LD-SimGIC, a groupwise semantic similarity measure for LD that adapts the intuition underlying SimGIC by replacing ontology annotations with incoming and outgoing RDF features weighted according to their information content.

2.2. Assessing Semantic Alignment in Knowledge Graph Embeddings

In recent years, an increasing number of studies [3, 4] have investigated whether KGEMs truly capture semantics, as so far this capability has been mainly tested for link prediction (LP) tasks, with purposely developed libraries (e.g. [17, 18]) and LP-based metrics. Specifically, Jain et al. [4] investigate whether embeddings in high-dimensional spaces are able to capture semantic similarity among entities belonging to the same class or concept. To this end, they perform a systematic analysis through both classification and clustering experiments on entity embeddings. Focusing on clustering, the underlying assumption is that, if embeddings effectively capture the features of entities, instances belonging to the same class should be grouped into the same cluster. Clustering performance, thus, is used as a proxy to assess the extent to which semantic information is preserved in the embedding space. Their results raise doubts about the expected spatial closeness of semantically similar entities, showing that current KGEMs have limited capability in capturing semantic features, and that their effectiveness is highly dependent on the characteristics of the dataset. Along the same line of investigation, Hubert et al. [3] analyze the extent to which entity similarity in KGs is reflected in the embedding space, i.e., whether semantic similarity is preserved as proximity in the embedding space. To this end, they compare the K-nearest neighbors obtained in the embedding space, based on cosine similarity, with those derived directly from the graph structure, where entity similarity is defined through the overlap of their local subgraphs (1-hop and 2-hop neighborhoods) measured using the Jaccard coefficient. Their results show that proximity in the embedding space does not consistently align with graph-based similarity, and that this relationship varies significantly across models and datasets, confirming the findings of Jain et al. [4]. Moreover, by correlating the degree of agreement between the two similarity notions with LP performance, they show that strong performance in LP does not necessarily translate into effectiveness in downstream tasks such as recommendation. Taken together, these findings motivate a direct investigation of the semantic information encoded by KG embeddings beyond their predictive performance. Building on this line of research, we assess whether geometric cohesion in the embedding space is associated with semantic similarity derived from the underlying KG. Although Jain et al. [4] also adopt a group-level perspective through clustering, their analysis focuses on the preservation of class-level information,

assessing whether entities belonging to the same class are grouped together. In contrast, our goal is to determine whether geometrically cohesive entities are also semantically similar according to the features explicitly represented in the KG.

Our work is closer in spirit to Hubert et al. [3], who explicitly compare similarity in the embedding space with similarity derived from the KG, focusing on resources within predefined entity classes. However, our study differs in both the notion of semantic similarity and the scope of the analysis. While Hubert et al. [3] measure graph-based similarity through the Jaccard overlap of local neighborhoods and focus on pairwise comparisons among the top-K nearest neighbors, we rely on information content-based semantic similarity for LD and examine the relationship between geometric and semantic similarity at both the pairwise and groupwise levels. We also consider a broader set of KGEMs and complement the quantitative correlation analysis with logic-based, natural-language explanations of the semantic commonalities shared by nearby resources. Finally, following the concerns raised by Jain et al. [4] about the semantic utility of KGEMs in complex real-world settings, we evaluate the models on two real-world KGs from different domains.

3. Methodology

This work investigates semantic alignment and interpretability in knowledge graph embeddings. Since proximity-based outputs, such as entity neighborhoods and clusters, are difficult to interpret when considered exclusively in the latent space, we relate their geometric cohesion to explicit semantic similarity derived from the underlying knowledge graph at both the pairwise and groupwise levels. In both settings, the association between geometric and semantic similarity is measured through correlation analysis. Finally, logic-based natural-language explanations are generated to make explicit the semantic features shared by geometrically close resources.

3.1. Pairwise Similarity

3.1.1. Geometric Cohesion

Cosine Similarity. Geometric similarity between two entities is measured by comparing their representations in the embedding space through cosine similarity. This measure captures the angular proximity between two entity embeddings, with values ranging from -1 to 1: 1 indicates identical directions, 0 indicates orthogonality, and -1 indicates opposite directions. Higher values therefore indicate that the corresponding entities are closer in the latent space. Cosine similarity is widely used in the KGEM literature [19, 3, 8], as it enables comparisons based on direction rather than scale. In our implementations, it is computed using the Scikit-learn *cosine_similarity* function [20].

3.1.2. Semantic Similarity

PICSS. The semantic similarity between pairs of resources is computed using PICSS [16]. Each resource is represented through a set of LD features extracted from its RDF descriptions. Each feature is assigned an IC value, which reflects how informative or specific the feature is within the dataset. Given an RDF triple $\langle s, p, o \rangle$, where s is the subject, p is the predicate, and o is the object, we consider as resources all entities that occur as subject or object and are identified by Uniform Resource Identifiers (URIs). Given an entity e, a feature f is defined as a triple $\langle p, e_t, D \rangle$, where p is the predicate connecting e to another entity e_t, and D indicates the direction of the relation, either *IN* or *OUT*. In the *IN* case, e occurs as the object of an RDF triple $\langle e_t, p, e \rangle$, while in the *OUT* case, e occurs as the subject of an RDF triple $\langle e, p, e_t \rangle$. Accordingly, incoming and outgoing features are defined as:

$$f_e^{\text{IN}} = \langle p, e_t, \mathit{IN} \rangle \quad \text{if } \langle e_t, p, e \rangle \in KG, \tag{1}$$

$$f_e^{\text{OUT}} = \langle p, e_t, \mathit{OUT} \rangle \quad \text{if } \langle e, p, e_t \rangle \in KG. \tag{2}$$

In both cases, e_t is required to be a resource identified by a URI. The feature set of an entity e is defined as the union of its outgoing (F_e^{OUT}) and incoming (F_e^{IN}) feature sets:

$$F_e = F_e^{\text{OUT}} \cup F_e^{\text{IN}}, \tag{3}$$

The information content of a feature f is defined as:

$$IC(f) = -\log_2 \frac{\phi(f)}{N}, \tag{4}$$

where $\phi(f)$ denotes the number of resources described by feature f, and N is the total number of resources in the KG.
Thus, given two entities e_i and e_j, PICSS is defined as:

$$\text{PICSS}(e_i, e_j) = \frac{\sum_{f \in F_{e_i} \cap F_{e_j}} IC(f)}{\sum_{f \in F_{e_i} \cap F_{e_j}} IC(f) + \sum_{f \in F_{e_i} \setminus F_{e_j}} IC(f) + \sum_{f \in F_{e_j} \setminus F_{e_i}} IC(f)}. \tag{5}$$

PICSS values range from 0 to 1, where 0 indicates that the two resources do not share any LD feature, and 1 indicates that their feature sets are identical with respect to the IC-weighted features considered.

3.2. Groupwise Similarity

3.2.1. Geometric Cohesion

Average Pairwise Cosine Similarity. The geometric cohesion of a cluster is measured as the average pairwise cosine similarity among all unordered pairs of resources within the cluster, consistent with the pairwise geometric similarity (Section 3.1). Therefore, its values range from -1 to 1. This measure summarizes how close, on average, the embeddings of the resources belonging to the same cluster are in the latent space. Higher values indicate clusters whose entities are more geometrically cohesive in the embedding space.

3.2.2. Semantic Similarity

LD-SimGIC. To assess semantic similarity at the groupwise level, we introduce Linked Data SimGIC, in brief **LD-SimGIC**, a novel groupwise semantic similarity measure for LD, to extend the pairwise setting addressed by PICSS to groups of resources. As introduced in Section 2, SimGIC [15] is an ontology-based groupwise similarity measure. Given a set of concepts, SimGIC compares their annotation sets, where the annotation set $A(c_i)$ denotes the set of ontology terms associated with concept c_i, by computing the ratio between the sum of the information content of the shared annotations and the sum of the information content of the annotations occurring in their union. Formally, given a group of concepts $C = \{c_1, \ldots, c_n\}$ and the annotation set $A(c_i)$ associated with each concept c_i, SimGIC can be expressed as:

$$\text{SimGIC}(C) = \frac{\sum_{a \in \bigcap_{c_i \in C} A(c_i)} IC(a)}{\sum_{a \in \bigcup_{c_i \in C} A(c_i)} IC(a)}. \tag{6}$$

Building on this idea, we define LD-SimGIC by replacing ontology annotations with LD features, as defined in Equations 1 and 2. Given a group of resources $E = \{e_1, \ldots, e_n\}$, where each resource e_i is represented by its feature set F_{e_i}, and each feature is weighted by its information content as defined in Equation 4, LD-SimGIC is defined as:

$$\text{LD-SimGIC}(E) = \frac{\sum_{f \in \bigcap_{e_i \in E} F(e_i)} IC(f)}{\sum_{f \in \bigcup_{e_i \in E} F(e_i)} IC(f)}. \tag{7}$$

In line with both PICSS and SimGIC, LD-SimGIC ranges from 0 to 1. When the intersection of the feature sets is empty, i.e., when no LD feature is shared by all resources in the group, the numerator is defined as zero and, consequently, LD-SimGIC is equal to 0. A value of 1 indicates maximum semantic similarity, i.e., that all resources share the same IC-weighted feature set.

3.3. Correlation Analysis

Kendall's Tau. The correlation between geometric and semantic similarity, in both the pairwise and groupwise settings, is assessed using Kendall's Tau[1]. This coefficient ranges from -1 to 1: positive values indicate increasing agreement between the geometric and semantic rankings, negative values indicate inverse agreement, and values close to 0 indicate little or no monotonic association. We adopt Kendall's Tau because our analysis focuses on rankings agreement rather than on the absolute similarity scores. In particular, we evaluate whether pairs or clusters ranked as more cohesive in the embedding space are also ranked as more semantically similar according to the corresponding LD similarity measure. For each model, we also report the associated p-value to assess statistical significance.

3.4. Explanation of Shared Features

(Least) Common Subsumer. The common characteristics of two or more RDF resources may be systematically derived using the notion of *Least Common Subsumer (LCS)*, originally defined for pairs of RDF resources [22] and later generalized to sets of resources [23], in order to capture cluster-wide commonalities.

The notion of LCS for RDF [24] is inspired by the corresponding concept in Description Logics [25], and is adapted to the specific features of RDF. More precisely, the LCS of a set of RDF resources[2] is defined as the most specific set of RDF triples, that may include blank nodes for generalizing different URIs, that is entailed[3] by every resource in that set.

From the RDF graphs that are compared to build the LCS, triples that match an optional set of user-defined *stop-patterns* (considered irrelevant) can be filtered out in advance — *e.g.,* drug packaging characteristics when the user is interested in their active ingredients. However, the resulting LCS often turns out to be still too fine-grained and overly specific, which reduces its usefulness for explaining commonalities — *e.g.,* it may say that each resource is related to some generic resource (a blank node) through some generic predicate (another blank node). For this reason, the LCS is post-processed by recursively removing all information that is irrelevant to explaining similarity, by specifying a set of triple patterns, called *uninformative triples*, that may depend from the application.

Following this double filtering, the explanation step relies on a (no longer "least") Common Subsumer (CS) that retains only the information considered relevant.

We note that theoretical results [26, 27] show that, for some worst-case sets of k resources, where each resource is described by a set of at most n triples, the size of the smallest logically equivalent representation of the LCS is exponential — specifically, $\Omega(n^k)$. However, such worst cases never appeared in our experiments, and the size of our CS was always manageable.

Natural Language Verbalization. A natural language explanation is generated from the CS representation in RDF through a Large Language Model (LLM), making human-readable the RDF features actually shared by nearby resources.

4. Experiments

4.1. Datasets

For the experiments, we consider two real-world knowledge graphs from different domains in order to assess the generality of our approach:

- *DrugBank* is a biomedical knowledge graph containing detailed information about drugs and their interactions with biological entities, such as targets, enzymes, carriers, and transporters [28].

[1] Relying on the *scipy.stats.kendalltau* function [21].
[2] Modelled as rooted graphs [24, Def. 5]
[3] According to Rooted Simple Entailment between rooted graphs [24, Def. 6]

Table 1
Statistics of the datasets in terms of number of triples, entities, and relations.

Dataset	Triples	Entities	Relations
DrugBank	3,649,531	1,522,943	95
SemOpenAlex	7,733,754	1,676,464	77

- *SemOpenAlex* is a scholarly knowledge graph containing over 26 billion RDF triples and describing publications, authors, institutions, and topics [29].

Due to the large scale of SemOpenAlex, we restrict our analysis to a subset of the graph; details on the extraction procedure are provided in *Appendix A*[4]. Main statistics of both datasets are reported in Table 1. From these datasets, we focus on drugs in DrugBank and works associated with the concept of *business* in SemOpenAlex, resulting in 7,678 and 4,747 resources, respectively. This choice is motivated by our goal of analyzing within-class semantic similarity, i.e., how similar drugs are to other drugs and business-related works to other works of the same type. The SPARQL queries used for their extraction are provided in *Appendix B* and *Appendix C*.

4.2. Knowledge Graph Embeddings

Knowledge graph embeddings are generated using the Python libraries PyKEEN [17] (version 1.11.1) and pyRDF2Vec [30] (version 0.2.3). A representative set of KGEMs is considered, namely *CrossE* [31], *DistMA* [8], *DistMult* [32], *ERMLP* [33], *HolE* [34], *MuRE* [35], *NTN* [36], *PairRE* [37], *ProjE* [38], *RESCAL* [39], *SE* [40], *TorusE* [41], *TransD* [42], *TransE* [43], *TransF* [44], *TransH* [45], *TransR* [46], *TuckER* [47], and *UM* [48], together with *pyRDF2Vec*. To ensure reproducibility, a random seed of 1 is fixed for all models. For PyKEEN-based models, this seed is used both for dataset splitting and embedding training. In contrast, for pyRDF2Vec, no train/validation/test split is required, as the model learns entity representations from walks extracted directly from the RDF graph; therefore, the seed is used only to make the embedding generation process reproducible. The entity embedding dimensionality is set to 100, as this is the default value in pyRDF2Vec and falls within the range adopted by PyKEEN for many models. It is also a common choice in the literature [37, 42, 36, 32], while representing a trade-off between model performance [42, 36] and computational cost, as both time and space complexity scale linearly with this hyperparameter [49]. For PyKEEN-based models, both DrugBank and SemOpenAlex are split into training, validation, and test sets with a 90%, 5%, and 5% ratio, respectively. Although PyKEEN supports evaluation through validation and testing, these phases are of limited relevance in our setting, as our goal is not to assess predictive performance but to obtain embeddings for the subsequent similarity and clustering analyses. For this reason, we prioritize maximizing the amount of training data, keeping validation and test sets intentionally small. Furthermore, we ensure that every entity appears at least once across the training, validation, and test sets, preventing unseen entities during embedding learning. For pyRDF2Vec, we set the main hyperparameters following the configuration suggested in [50, 3], in addition to the random seed and vector size defined above. In particular, the maximum number of walks per entity is set to 2,000, the walk depth to 4, the window size to 5, and the underlying Word2Vec model is trained for 25 epochs using the skip-gram architecture. All other unspecified hyperparameters for both PyKEEN-based models and pyRDF2Vec are kept at their default values. All experiments in this work, including embedding generation, are conducted on a machine equipped with a 12th Gen Intel Core i7-12700F processor (2.10GHz, 12 cores), an NVIDIA GeForce RTX 4060 Ti GPU, 64 GB of RAM, and a 2TB Kingston SNV2S2000G SSD.

4.3. Pairwise similarity

To evaluate whether semantic similarity is reflected at the level of individual resource pairs, we consider all possible unordered pairs of the selected resources in each dataset, i.e., all pairwise combinations

[4]All appendices are provided in Supplementary Material (https://doi.org/10.5281/zenodo.22810772)

excluding self-pairs and duplicate symmetric pairs. Given the 7,678 drugs extracted from DrugBank and the 4,747 works extracted from SemOpenAlex (as described in Section 4.1) this results in 29,472,003 and 11,264,631 resource pairs, respectively. For each pair, we compute pairwise geometric and semantic similarity scores as described in Section 3.1.

4.4. Groupwise similarity

To assess semantic similarity at a broader level than individual entity pairs, we use clustering as a means to group the selected resources — drugs in DrugBank and works in SemOpenAlex, as defined in Section 4.1 — that are close to one another in the embedding space. The goal is not to optimize or compare clustering algorithms, but rather to obtain geometrically coherent groups of resources and evaluate whether such groups are also semantically coherent according to LD features. Clustering is performed independently for each KGEM and dataset; therefore, partitions may differ across models, while correlation is assessed within each model between geometric cohesion (as described in Section 3.2) and LD-SimGIC. Specifically, for each KGEM, Hierarchical Agglomerative Clustering is applied to the entity vectors generated by that model, specifying only the *distance metric*, the *linkage criterion*, and the *number of clusters* (k), while keeping all other hyperparameters at their default values. In particular, *cosine distance* is used as distance metric consistently with the geometric cohesion definitions introduced in Sections 3.1 and 3.2. Embeddings are L2-normalized prior to clustering to place them on the unit hypersphere, aligning the Euclidean geometry underlying the Calinski–Harabasz criterion with the angular geometry adopted for clustering. *Average linkage* is adopted as it considers all pairwise distances between elements of the two clusters, making it less sensitive to outliers compared to single and complete linkage, which rely on the closest and farthest pair of points, respectively. Ward linkage is not considered, as it is not compatible with cosine distance. For each KGEM and dataset, k is selected independently using the Calinski–Harabasz index, choosing the value that maximizes it over a predefined range (i.e., [80,570] for DrugBank and [50,350] for SemOpenAlex). These ranges ensure sufficient coverage of the solution space while avoiding excessively small or large clusterings. Assuming roughly uniform clusters, this yields an average cluster size of approximately 13-95 resources. All clustering and evaluation procedures are implemented using the Scikit-learn library[5] [20].

4.5. Explanation of Shared Features

Common Subsumer. For each model, the CS is computed up to depth 4 for the most geometrically cohesive pairs and clusters, as a qualitative inspection of the strongest geometric proximities. During this step, irrelevant information is filtered out from the resource KGs by using dataset-specific *stop-patterns*. Examples include identifiers, labels, and generic type assertions in DrugBank, as well as DOIs and citation counts in SemOpenAlex. The complete *stop-patterns* and *uninformative triples* used for DrugBank and SemOpenAlex are reported in *Appendix D* and *Appendix E*, respectively.

Natural Language Verbalization. Gemini[6] is used to verbalize the CSs from RDF into natural language. We report the prompt and instructions in *Appendix F*.

5. Results and Discussion

5.1. Pairwise Similarity

The pairwise results (Table 2) show an overall lack of meaningful association between geometric cohesion and semantic similarity. Although Kendall's Tau is statistically significant for many models under the conventional threshold of $p < 0.05$, its magnitude is generally close to zero. Thus, cosine similarity does not reliably induce a ranking of resource pairs consistent with their PICSS semantic

[5]Relying on the *AgglomerativeClustering* and *calinski_harabasz_score* functions.
[6]Version 3 Flash Preview, with all hyperparameters left at their default values.

Table 2
Pairwise correlation analysis between cosine similarity and PICSS for DrugBank and SemOpenAlex. Statistically significant p-values ($p < 0.05$) are reported in italic. Kendall's Tau columns use a white-to-dark-green scale from the most negative to the strongest positive correlation within each column.

	DrugBank		SemOpenAlex	
Model	**Kendall's Tau**	**p-value**	**Kendall's Tau**	**p-value**
CrossE	-0.05717	*0.00000*	0.01738	*0.00000*
DistMA	-0.02522	*0.00000*	-0.00833	*0.00000*
DistMult	-0.19332	*0.00000*	0.07552	*0.00000*
ERMLP	-0.02581	*0.00000*	0.00333	*0.00000*
HolE	-0.10588	*0.00000*	0.02022	*0.00000*
MuRE	-0.00997	*0.00000*	0.07868	*0.00000*
NTN	0.00015	0.23521	0.00033	0.09223
PairRE	0.00097	*0.00000*	0.00879	*0.00000*
ProjE	-0.12860	*0.00000*	0.01093	*0.00000*
pyRDF2Vec	0.17907	*0.00000*	0.16439	*0.00000*
RESCAL	-0.12650	*0.00000*	0.00001	0.96609
SE	-0.00262	*0.00000*	0.02188	*0.00000*
TorusE	0.00138	*0.00000*	-0.00010	0.61301
TransD	-0.15411	*0.00000*	0.04387	*0.00000*
TransE	-0.08755	*0.00000*	-0.03590	*0.00000*
TransF	-0.02066	*0.00000*	0.00002	0.92582
TransH	-0.18816	*0.00000*	0.01145	*0.00000*
TransR	-0.04748	*0.00000*	0.00260	*0.00000*
TuckER	-0.14420	*0.00000*	-0.03017	*0.00000*
UM	0.00466	*0.00000*	0.02385	*0.00000*

similarity. This behavior holds for all models except NTN on DrugBank and NTN, RESCAL, TorusE, and TransF on SemOpenAlex, for which Kendall's Tau differs from zero by at most 0.00033 and the corresponding p-values are above the significance threshold. Hence, in these cases, there is no statistical evidence of a monotonic association between cosine similarity and PICSS. Overall, Kendall's Tau remains weak across datasets; even when statistically significant, it does not indicate a strong or consistent monotonic relationship between geometric proximity in the embedding space and semantic similarity.

A possible concern is that these results are driven by the large number of pairwise comparisons considered in the analysis (see Section 4.3). To further support our findings, we report in *Appendix G* the results of the same correlation analysis restricted to the top-1,000 most geometrically similar pairs for each model. The results of this restricted analysis are consistent with the full setting: Kendall's Tau remains weak, with the largest absolute deviation from zero equal to 0.16449 for pyRDF2Vec. When p-values exceed 0.05, there is no statistical evidence of monotonic correlation; when they are below 0.05, Kendall's Tau is still too small to support a meaningful association between geometric cohesion and semantic similarity.

5.2. Groupwise Similarity

The groupwise results (Table 3) show a different trend from the pairwise analysis. In this setting, geometric cohesion appears to be meaningfully associated with semantic similarity for several KGEMs.

Specifically, many models exhibit positive and statistically significant Kendall's Tau values, indicating that more cohesive groups in the embedding space also tend to be more semantically coherent according to LD-SimGIC. For DrugBank, the strongest correlations are observed for NTN, TransF, TorusE, UM, and DistMA, all showing Kendall's Tau values above 0.70. Other models, such as CrossE, PairRE, ProjE, and SE, also exhibit relatively high positive correlations, with values above 0.60. These results

Table 3
Groupwise correlation analysis between cosine similarity and LD-SimGIC for DrugBank and SemOpenAlex. Statistically significant p-values ($p < 0.05$) are reported in italic. Kendall's Tau columns use a white-to-dark-green scale from the most negative to the strongest positive correlation within each column.

	DrugBank		SemOpenAlex	
Model	**Kendall's Tau**	**p-value**	**Kendall's Tau**	**p-value**
CrossE	0.62714	*0.00000*	0.78047	*0.00000*
DistMA	0.70886	*0.00000*	0.81714	*0.00000*
DistMult	0.12381	0.55897	-0.25275	0.23311
ERMLP	0.39095	*0.00000*	0.38736	*0.00000*
HolE	0.08298	*0.04537*	0.47611	*0.00000*
MuRE	0.25603	*0.00000*	0.44044	*0.00000*
NTN	0.78608	*0.00000*	0.79429	*0.00000*
PairRE	0.67532	*0.00000*	0.78939	*0.00000*
ProjE	0.62769	*0.00000*	0.77166	*0.00000*
pyRDF2Vec	0.40380	*0.00000*	-0.03723	0.70893
RESCAL	-0.00686	0.91150	0.68000	*0.00000*
SE	0.67722	*0.00000*	0.79316	*0.00000*
TorusE	0.76519	*0.00000*	0.67673	*0.00000*
TransD	0.25536	*0.00039*	0.04413	0.32341
TransE	0.13460	0.07346	0.47751	*0.00000*
TransF	0.76772	*0.00000*	0.77143	*0.00000*
TransH	-0.33333	0.25952	-0.20468	0.23778
TransR	0.06671	0.37200	0.06286	0.51951
TuckER	0.31061	*0.00003*	-0.23529	0.20490
UM	0.74367	*0.00000*	0.79099	*0.00000*

suggest that, for these models, geometric cohesion captures a meaningful degree of group-level semantic coherence. However, this behavior is not consistent across all models. DistMult, RESCAL, TransE, TransH, and TransR do not show statistically significant correlations (p>0.05), while HolE exhibits only a weak positive association despite being statistically significant. A similar pattern can be observed for SemOpenAlex: CrossE, DistMA, NTN, PairRE, ProjE, SE, TransF, and UM achieve strong positive and statistically significant correlations, with Kendall's Tau values around or above 0.77. This indicates a strong monotonic association between geometric cohesion and LD-SimGIC for these models. Statistically significant correlations are also observed for ERMLP, HolE, MuRE, RESCAL, TorusE, and TransE, although with lower magnitudes. Conversely, DistMult, TransD, TransH, TransR, and TuckER do not provide statistical evidence of a monotonic association, as their p-values are above 0.05.

5.3. Explanation of Shared Features

This section provides natural language explanations of the commonalities shared by resources grouped within the same clusters in order to support a qualitative evaluation complementary to the quantitative analysis. We focus on clusters rather than on individual pairs because the main evidence of semantic preservation emerges at the groupwise level. Specifically, across the two datasets, we report the explanation of the most cohesive cluster for the model achieving the highest correlation score in Table 3, namely DistMA on SemOpenAlex with a Kendall's Tau of 0.81714, as well as that of the most cohesive cluster for the model achieving the lowest correlation score, namely TransH on DrugBank with a Kendall's Tau of -0.33333. The most cohesive cluster produced by DistMA on SemOpenAlex consists of 28 resources. Its logic-based explanation in natural language is reported below[7]:

[7]Its RDF/Turtle representation is provided in *Appendix H*.

The set of resources identified as this work is characterized by a specific creator, a title, and a publication year. This work, which is classified by a particular work type and crossref type, is not a paratext and has not been retracted. The work features an authorship at position zero that includes a specific author and a corresponding status indicator. It is associated with a primary location, which also serves as its general location; this location is identified by a specific URL and a certain open access status. The work's broader open access properties further detail its status and whether any repository contains the full text. Additionally, the work is linked to several concepts. One concept is related to economics and another specific resource, while a second concept connects the work to engineering, economics, political science, philosophy, and mathematics. A third concept is associated with a specific term, and the work is further related to the concept of sociology. Sociology, in turn, is linked to a wide range of fields including physics, business, biology, engineering, computer science, geography, economics, psychology, history, political science, art, philosophy, mathematics, chemistry, medicine, law, and materials science. Finally, the work includes several specific internal properties and boolean attributes.

For TransH on DrugBank, the most cohesive cluster consists of 7,563 resources out of the 7,678 drugs considered in the analysis. However, its CS contains no triples, indicating that the resources in this cluster do not share any informative RDF-based commonality beyond belonging to the same general class, i.e., being drugs. To further inspect this behavior, we also analyze the second most cohesive cluster produced by TransH. This cluster consists of 5 resources and yields the following shared information[8]:

The set of resources is categorized as being of the type small molecule.

Some observations may be drawn from the results above:

- No statistically significant correlation emerges between geometric cohesion and semantic similarity (PICSS) in the pairwise comparison of resources in both DrugBank and SemOpenAlex. This result is consistent with the observations by Hubert et al. [3], who, by comparing pairs of nearest-neighbour resources, show that entities that are close in the embedding space are not necessarily similar according to graph-based notions of semantic similarity. Our findings further support this claim by showing that the lack of alignment also holds when semantic similarity is assessed through a measure grounded in information-theoretic principles, rather than only through feature-overlap-based approaches. From an interpretability perspective, this suggests that geometric proximity alone is insufficient to provide a semantic characterization of the similarity between two resources.
- When reverting to groupwise similarity, 5 KGEMs on DrugBank and 8 on SemOpenAlex show a significant correlation between geometric and semantic similarity (LD-SimGIC); this behaviour suggests a greater ability of such models to retain the informative content of resources grouped by a clustering process. This represents one of the main contributions of our analysis, since previous studies have mostly focused on pairwise evaluations, whereas here the correlation is assessed over groups of resources considered simultaneously. Although these models may be limited for tasks such as recommendation or for representing individual entities through isolated vectors, the results suggest that some of them are nevertheless able to capture a form of global semantic structure. At the same time, our findings remain aligned with Jain et al. [4] and Hubert et al. [3]'s observation that embedding performance depends not only on the model itself, but also on the dataset under consideration. Indeed, several models, such as DistMult, HolE, pyRDF2Vec, RESCAL, TransE, and TuckER exhibit noticeable differences in the correlation values observed across DrugBank and SemOpenAlex.
- The models with high correlation between cosine similarity and LD-SimGIC seem to produce clusters with a significant number of commonalities, according to the implemented logic-based explanation process. This suggests that, when a KGEM preserves groupwise semantic similarity, the resulting clusters are not only geometrically coherent but also more interpretable from a symbolic and KG-based perspective.

[8] Its RDF/Turtle representation is provided in *Appendix I*.

6. Conclusion

In this paper, we investigated whether geometric cohesion in knowledge graph embedding spaces reflects semantic similarity in KGs. This question is relevant to the interpretation of proximity-based outputs, since nearby entity representations and embedding-induced clusters are difficult to inspect when considered only in the latent space. We addressed this question at both the pairwise and groupwise levels, comparing cosine similarity with PICSS for entity pairs and introducing LD-SimGIC, an IC-based measure for assessing the semantic cohesion of groups of RDF resources. The evaluation was conducted on two real-world KGs, DrugBank and SemOpenAlex, covering biomedical and scholarly domains. The results show a clear distinction between the two settings: pairwise correlations are generally very weak, while stronger and statistically significant correlations emerge at the groupwise level for several KGEMs, most consistently CrossE, DistMA, NTN, PairRE, ProjE, SE, TransF, and UM. This suggests that, although embeddings may not reliably capture fine-grained semantic similarity between individual entities, some models can capture meaningful semantic structure at the level of resource groups. However, this behavior is not uniform across KGEMs and datasets, highlighting that the alignment between geometric and semantic cohesion is both dataset- and model-dependent. Finally, we complemented the quantitative analysis with a qualitative analysis based on logic-based, natural language explanations of cluster commonalities. These explanations help clarify what geometrically cohesive clusters actually share from a semantic perspective, showing that models with stronger groupwise correlations tend to produce more meaningful commonalities. Future work may further investigate the correlation between LD-SimGIC and quantitative measures directly derived from cluster explainability, as well as explore refining entity embeddings based on the specificity of their corresponding KG nodes, to assess whether this leads to more meaningful geometric similarity.

Acknowledgments

This work is co-funded by the Italian aerospace technological district “Distretto Tecnologico Aerospaziale S.C.a.r.l.”

Declaration on Generative AI

Generative AI is used exclusively as an experimental component for the natural-language verbalization of RDF Common Subsumers, as described in Sections 3.4 and 4.5. It is not used to generate the scientific content, results, or conclusions of this manuscript.

References

[1] A. Rossi, D. Barbosa, D. Firmani, A. Matinata, P. Merialdo, Knowledge graph embedding for link prediction: A comparative analysis, ACM Trans. Knowl. Discov. Data 15 (2021). doi:10.1145/3424672.

[2] A. Domingues, N. Jain, A. M. Peñuela, E. Simperl, Bringing Back Semantics to Knowledge Graph Embeddings: An Interpretability Approach, in: T. R. Besold, A. d’Avila Garcez, E. Jimenez-Ruiz, R. Confalonieri, P. Madhyastha, B. Wagner (Eds.), Neural-Symbolic Learning and Reasoning, Springer Nature Switzerland, Cham, 2024, pp. 192–203.

[3] N. Hubert, H. Paulheim, A. Brun, D. Monticolo, Do similar entities have similar embeddings?, in: The Semantic Web: 21st International Conference, ESWC 2024, Hersonissos, Crete, Greece, May 26–30, 2024, Proceedings, Part I, Springer-Verlag, Berlin, Heidelberg, 2024, p. 3–21. doi:10.1007/978-3-031-60626-7_1.

[4] N. Jain, J.-C. Kalo, W.-T. Balke, R. Krestel, Do embeddings actually capture knowledge graph semantics?, in: The Semantic Web: 18th International Conference, ESWC 2021, Virtual Event,

June 6–10, 2021, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2021, p. 143–159. doi:10.1007/978-3-030-77385-4_9.

[5] P. Szmeja, M. Ganzha, M. Paprzycki, W. Pawłowski, Dimensions of Semantic Similarity, Springer International Publishing, Cham, 2018, pp. 87–125. doi:10.1007/978-3-319-67946-4_3.

[6] R. T. Sousa, S. Silva, C. Pesquita, Explaining protein–protein interactions with knowledge graph-based semantic similarity, Computers in Biology and Medicine 170 (2024) 108076. doi:10.1016/j.compbiomed.2024.108076.

[7] Q. Guo, F. Zhuang, C. Qin, H. Zhu, X. Xie, H. Xiong, Q. He, A survey on knowledge graph-based recommender systems, IEEE Transactions on Knowledge and Data Engineering 34 (2022) 3549–3568. doi:10.1109/TKDE.2020.3028705.

[8] X. Shi, Y. Xiao, Modeling multi-mapping relations for precise cross-lingual entity alignment, in: K. Inui, J. Jiang, V. Ng, X. Wan (Eds.), Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, Hong Kong, China, 2019, pp. 813–822. doi:10.18653/v1/D19-1075.

[9] S. Colucci, F. M. Donini, V. Schena, F. Scioscia, E. Di Sciascio, LISE: A Logic-Based Interactive Similarity Explainer for Clusters of RDF Data, IEEE Access 13 (2025) 90109–90128. doi:10.1109/ACCESS.2025.3571518.

[10] V. Schena, S. Colucci, F. Scioscia, F. M. Donini, Human-in-the-loop refinement of knowledge graph embeddings for semantic similarity, in: L. M. Camarinha-Matos, F. Ferrada (Eds.), Technological Innovation to Tackle Societal Challenges, Springer Nature Switzerland, Cham, 2026, pp. 358–369.

[11] P. Resnik, Using information content to evaluate semantic similarity in a taxonomy, in: Proceedings of the 14th International Joint Conference on Artificial Intelligence-Volume 1, 1995, pp. 448–453.

[12] D. Lin, An information-theoretic definition of similarity, in: Proceedings of the Fifteenth International Conference on Machine Learning, 1998, pp. 296–304.

[13] A. Adhikari, B. Dutta, A. Dutta, Finding most informative common ancestor in cross-ontological semantic similarity assessment: An intrinsic information content-based approach, Expert Systems with Applications 192 (2022) 116281. doi:https://doi.org/10.1016/j.eswa.2021.116281.

[14] C. Pesquita, D. Faria, A. O. Falcão, P. Lord, F. M. Couto, Semantic similarity in biomedical ontologies, PLOS Computational Biology 5 (2009) 1–12. doi:10.1371/journal.pcbi.1000443.

[15] C. Pesquita, D. Faria, H. Bastos, A. Falco, F. Couto, Evaluating go-based semantic similarity measures, Proc 10th Annual Bio-Ontologies Meeting (2007).

[16] R. Meymandpour, J. G. Davis, A semantic similarity measure for linked data: An information content-based approach, Knowledge-Based Systems 109 (2016) 276–293. doi:10.1016/j.knosys.2016.07.012.

[17] M. Ali, M. Berrendorf, C. T. Hoyt, L. Vermue, S. Sharifzadeh, V. Tresp, J. Lehmann, PyKEEN 1.0: A Python Library for Training and Evaluating Knowledge Graph Embeddings, J. Mach. Learn. Res. 22 (2021).

[18] X. Han, S. Cao, X. Lv, Y. Lin, Z. Liu, M. Sun, J. Li, OpenKE: An open toolkit for knowledge embedding, in: E. Blanco, W. Lu (Eds.), Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Association for Computational Linguistics, Brussels, Belgium, 2018, pp. 139–144. doi:10.18653/v1/D18-2024.

[19] C. d'Amato, P. Masella, N. Fanizzi, An approach based on semantic similarity to explaining link predictions on knowledge graphs, in: IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT '21, Association for Computing Machinery, New York, NY, USA, 2022, p. 170–177. URL: https://doi.org/10.1145/3486622.3493956. doi:10.1145/3486622.3493956.

[20] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.

[21] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov,

A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0 Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods 17 (2020) 261–272. doi:10.1038/s41592-019-0686-2.

[22] S. Colucci, F. M. Donini, E. Di Sciascio, Common subsumbers in RDF, in: M. Baldoni, C. Baroglio, G. Boella, R. Micalizio (Eds.), AI*IA 2013: Advances in Artificial Intelligence, Springer International Publishing, Cham, 2013, pp. 348–359.

[23] S. Colucci, F. M. Donini, E. Di Sciascio, Computing the Commonalities of Clusters in Resource Description Framework: Computational Aspects, Data 9 (2024). doi:10.3390/data9100121.

[24] S. Colucci, F. Donini, S. Giannini, E. Di Sciascio, Defining and computing Least Common Subsumers in RDF, Web Semant. 39 (2016) 62–80. doi:10.1016/j.websem.2016.02.001.

[25] W. W. Cohen, A. Borgida, H. Hirsh, Computing Least Common Subsumers in Description Logics, in: W. R. Swartout (Ed.), Proceedings of the 10th National Conference on Artificial Intelligence, San Jose, CA, USA, July 12-16, 1992, AAAI Press / The MIT Press, 1992, pp. 754–760.

[26] F. Baader, R. Küsters, R. Molitor, Computing Least Common Subsumer in Description Logics with Existential Restrictions, in: Proceedings of the 16th International Joint Conference on Artifical Intelligence-Volume 1, 1999, pp. 96–101. doi:10.25368/2022.85.

[27] G. Amendola, M. Manna, A. Ricioppo, A logic-based framework for characterizing nexus of similarity within knowledge bases, Information Sciences 664 (2024). doi:10.1016/j.ins.2024.120331.

[28] D. S. Wishart, C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, J. Woolsey, DrugBank: a comprehensive resource for in silico drug discovery and exploration, Nucleic Acids Research 34 (2006) D668–D672.

[29] M. Färber, D. Lamprecht, J. Krause, L. Aung, P. Haase, SemOpenAlex: The Scientific Landscape in 26 Billion RDF Triples, in: The Semantic Web – ISWC 2023: 22nd International Semantic Web Conference, Athens, Greece, November 6–10, 2023, Proceedings, Part II, Springer-Verlag, Berlin, Heidelberg, 2023, p. 94–112.

[30] B. Steenwinckel, G. Vandewiele, T. Agozzino, F. Ongenae, pyRDF2Vec: A Python Implementation and Extension of RDF2Vec, in: The Semantic Web: 20th International Conference, ESWC 2023, Hersonissos, Crete, Greece, May 28–June 1, 2023, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2023, p. 471–483. doi:10.1007/978-3-031-33455-9_28.

[31] W. Zhang, B. Paudel, W. Zhang, A. Bernstein, H. Chen, Interaction embeddings for prediction and explanation in knowledge graphs, in: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, WSDM '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 96–104. URL: https://doi.org/10.1145/3289600.3291014. doi:10.1145/3289600.3291014.

[32] B. Yang, W. tau Yih, X. He, J. Gao, L. Deng, Embedding entities and relations for learning and inference in knowledge bases, in: International Conference on Learning Representations, 2014.

[33] X. Dong, E. Gabrilovich, G. Heitz, W. Horn, N. Lao, K. Murphy, T. Strohmann, S. Sun, W. Zhang, Knowledge vault: a web-scale approach to probabilistic knowledge fusion, in: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14, Association for Computing Machinery, New York, NY, USA, 2014, p. 601–610. URL: https://doi.org/10.1145/2623330.2623623. doi:10.1145/2623330.2623623.

[34] M. Nickel, L. Rosasco, T. Poggio, Holographic embeddings of knowledge graphs, in: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, AAAI'16, AAAI Press, 2016, p. 1955–1961.

[35] I. Balažević, C. Allen, T. Hospedales, Multi-relational poincaré graph embeddings, volume 32, Curran Associates Inc., Red Hook, NY, USA, 2019. URL: https://proceedings.neurips.cc/paper_files/paper/2019/file/f8b932c70d0b2e6bf071729a4fa68dfc-Paper.pdf.

[36] R. Socher, D. Chen, C. D. Manning, A. Y. Ng, Reasoning with neural tensor networks for knowledge base completion, in: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 1, NIPS'13, Curran Associates Inc., Red Hook, NY, USA, 2013, p.

926–934.

[37] L. Chao, J. He, T. Wang, W. Chu, PairRE: Knowledge graph embeddings via paired relation vectors, in: C. Zong, F. Xia, W. Li, R. Navigli (Eds.), Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Association for Computational Linguistics, Online, 2021, pp. 4360–4369. doi:10.18653/v1/2021.acl-long.336.

[38] B. Shi, T. Weninger, Proje: embedding projection for knowledge graph completion, in: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, AAAI'17, AAAI Press, 2017, p. 1236–1242.

[39] M. Nickel, V. Tresp, H.-P. Kriegel, A three-way model for collective learning on multi-relational data, in: Proceedings of the 28th International Conference on International Conference on Machine Learning, ICML'11, Omnipress, Madison, WI, USA, 2011, p. 809–816.

[40] A. Bordes, J. Weston, R. Collobert, Y. Bengio, Learning structured embeddings of knowledge bases, Proceedings of the AAAI Conference on Artificial Intelligence 25 (2011) 301–306. URL: https://ojs.aaai.org/index.php/AAAI/article/view/7917. doi:10.1609/aaai.v25i1.7917.

[41] T. Ebisu, R. Ichise, Toruse: knowledge graph embedding on a lie group, in: Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence, AAAI'18/IAAI'18/EAAI'18, AAAI Press, 2018.

[42] G. Ji, S. He, L. Xu, K. Liu, J. Zhao, Knowledge Graph Embedding via Dynamic Mapping Matrix, in: C. Zong, M. Strube (Eds.), Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Association for Computational Linguistics, Beijing, China, 2015, pp. 687–696. doi:10.3115/v1/P15-1067.

[43] A. Bordes, N. Usunier, A. Garcia-Durán, J. Weston, O. Yakhnenko, Translating embeddings for modeling multi-relational data, in: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, NIPS'13, Curran Associates Inc., Red Hook, NY, USA, 2013, p. 2787–2795.

[44] J. Feng, M. Huang, M. Wang, M. Zhou, Y. Hao, X. Zhu, Knowledge graph embedding by flexible translation, in: Proceedings of the Fifteenth International Conference on Principles of Knowledge Representation and Reasoning, KR'16, AAAI Press, 2016, p. 557–560.

[45] Z. Wang, J. Zhang, J. Feng, Z. Chen, Knowledge graph embedding by translating on hyperplanes, in: AAAI Conference on Artificial Intelligence, volume 28, 2014. URL: https://ojs.aaai.org/index.php/AAAI/article/view/8870. doi:10.1609/aaai.v28i1.8870.

[46] Y. Lin, Z. Liu, M. Sun, Y. Liu, X. Zhu, Learning entity and relation embeddings for knowledge graph completion, in: AAAI Conference on Artificial Intelligence, volume 29, 2015. URL: https://ojs.aaai.org/index.php/AAAI/article/view/9491. doi:10.1609/aaai.v29i1.9491.

[47] I. Balazevic, C. Allen, T. Hospedales, TuckER: Tensor factorization for knowledge graph completion, in: K. Inui, J. Jiang, V. Ng, X. Wan (Eds.), Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, Hong Kong, China, 2019, pp. 5185–5194. URL: https://aclanthology.org/D19-1522/. doi:10.18653/v1/D19-1522.

[48] A. Bordes, X. Glorot, J. Weston, Y. Bengio, A semantic matching energy function for learning with multi-relational data, Mach. Learn. 94 (2014) 233–259. URL: https://doi.org/10.1007/s10994-013-5363-6. doi:10.1007/s10994-013-5363-6.

[49] J. Cao, J. Fang, Z. Meng, S. Liang, Knowledge graph embedding: A survey from the perspective of representation spaces, ACM Comput. Surv. 56 (2024). doi:10.1145/3643806.

[50] J. Portisch, N. Heist, H. Paulheim, Knowledge graph embedding for data mining vs. knowledge graph embedding for link prediction – two sides of the same coin?, Semant. Web 13 (2022) 399–422. doi:10.3233/SW-212892.