



# City Research Online

## City St George's, University of London

**Citation:** Herron, D., Jiménez-Ruiz, E. & Weyde, T. (2025). On the Potential of Logic and Reasoning in Neurosymbolic Systems using OWL-based Knowledge Graphs. *Neurosymbolic Artificial Intelligence*, 1, 29498732251320043. doi: 10.1177/29498732251320043

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

**Permanent repository link:** <https://openaccess.city.ac.uk/id/eprint/32688/>

**Link to published version:** <https://doi.org/10.1177/29498732251320043>

**Copyright and Reuse:** Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

# On the Potential of Logic and Reasoning in Neurosymbolic Systems Using OWL-Based Knowledge Graphs

Neurosymbolic Artificial Intelligence

Volume 1: 1–15

© The Author(s) 2025

Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/29498732251320043

[journals.sagepub.com/home/nai](https://journals.sagepub.com/home/nai)

David Herron , Ernesto Jiménez-Ruiz  and Tillman Weyde 

**Editor:** Roberto Confalonieri, University of Padova, Italy

**Solicited reviews:** Janna Hastings, University of Zurich, Switzerland; one anonymous reviewer

## Abstract

Knowledge graphs (KGs) feature ever more frequently as symbolic components in neurosymbolic research and systems. But even though a central concern of neurosymbolic artificial intelligence is to combine neural learning with symbolic reasoning, relatively little neurosymbolic research focuses on leveraging the logical representation and reasoning capabilities of Web Ontology Language (OWL)-based KGs. The objective of this position article is to inspire more neurosymbolic researchers to embrace the OWL and the Semantic Web by raising awareness of the benefits, capabilities, and applications of OWL-based KGs, particularly with respect to logical reasoning. We describe the ecosystem of open W3C standards-based resources available that support the adoption and use of OWL-based KGs; we describe tools that exist for engineering custom OWL ontologies tailored to particular research needs; we discuss the encoding of background KG knowledge in subsymbolic embedding spaces and various applications of this approach; we discuss and illustrate the reasoning capabilities of OWL-based KGs; and we describe several promising directions for research that focus on leveraging these reasoning capabilities. We also discuss the specialised resources needed to undertake research on OWL-based KGs in neurosymbolic systems. We use the example of NeSy4VRD, an image dataset with a custom-designed companion OWL ontology. The scarcity of this kind of resource should be addressed to accelerate research in this field.

## Keywords

neurosymbolic, artificial intelligence, deep learning, Semantic Web, OWL, ontologies, knowledge graphs, reasoning

Received: 29 July 2023; accepted: 04 November 2024

## 1 Introduction

Following a long gestation spanning decades, neurosymbolic artificial intelligence (NeSy AI) has recently blossomed into a recognised subfield of AI. While neural and symbolic traditions of AI have been tribally rival, recently there is a vibrant diversity of approaches blending the two (Sarker et al., 2021). Prompted by analysis of the limitations of deep learning (in, e.g. Chollet, 2018; Kautz, 2022; Marcus, 2018a, 2018b, 2020), and despite the recent advances resulting from scaling up deep learning, as evidenced in large language models (LLMs), increasing number of researchers are drawn to NeSy AI. The shared motivation is to explore combinations of neural learning and symbolic knowledge representations in order to get the best of both worlds, in a shared belief that this is the best route for advancing AI towards artificial general intelligence.

Knowledge graphs (KGs) are representations of symbolic knowledge that conform to a graph model, where nodes are concepts and entities of interest, and edges are relationships between them (Hogan et al., 2021; Kejriwal et al., 2021).

Department of Computer Science, City St George's, University of London, London, UK

### Corresponding Author:

David Herron, Department of Computer Science, City St George's, University of London, Northampton Square, London EC1V 0HB, UK.

Email: [david.herron@city.ac.uk](mailto:david.herron@city.ac.uk)



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further

permission provided the original work is attributed as specified on the SAGE and Open Access page

(<https://uk.sagepub.com/en-gb/eur/open-access-at-sage>).

As NeSy research has expanded, so has the frequency with which KGs feature as symbolic components in hybrid, NeSy systems (Hitzler, 2021). One example of this is the progressively developing theme of ‘deep deductive reasoning’ (Bianchi & Hitzler, 2019; Ebrahimi et al., 2021a, 2021b), where neural networks (NNs) are trained to reason over KGs. KGs have also been shown to be helpful when data samples are expensive, difficult or impossible to obtain, such that there is a lack of data with which to train robust deep learning-based systems, as in few-shot and zero-shot learning scenarios (Chen & Chen, 2021; Chen & Geng, 2021; Geng et al., 2021).

The Web Ontology Language (OWL) (Hitzler et al., 2012; Motik et al., 2012) is a key component of the Semantic Web technology stack (The Semantic Web Stack, 2022; The Semantic Web Wiki, 2001). OWL ontologies (semantic schemas enriched with logic that describe domain knowledge symbolically) govern Semantic Web (SW) KGs by specifying what assertions of knowledge (types of triples) are admissible and inadmissible. Inference semantics (ontological rules) associated with OWL constructs permit reasoning algorithms to reason over OWL ontologies (and associated KG data, vast or tiny), both to infer new knowledge (new triples) and to enforce logical consistency constraints. With suitable ontology design, inference can be used as prediction, and incremental (on-the-fly) reasoning can facilitate real-time interaction. In summary, OWL-based KGs can be used as symbolic deduction engines in NeSy systems. The NeSy system AlphaGeometry (Trinh et al., 2024) combines an LLM with a symbolic deduction engine (Horn clause geometry and algebra rules, plus inference algorithms) to solve geometry problems. OWL-based KG technologies offer researchers the option to explore combining neural learning and symbolic reasoning in ways analogous to AlphaGeometry, by using OWL-based symbolic deduction during NN training and/or inference.

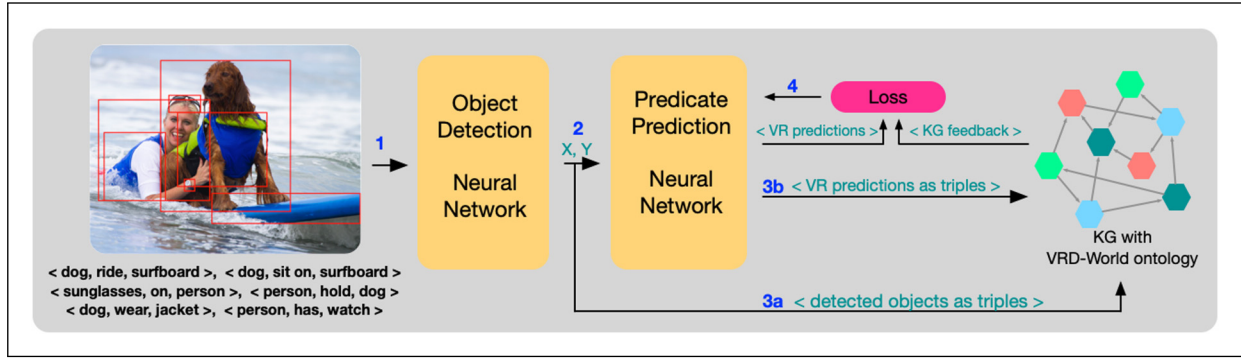
Given that combining neural learning with symbolic reasoning is central to NeSy AI, it is surprising how scant the literature is that explores applications of OWL reasoning in NeSy systems. A systematic mapping study of 476 recent papers that combine Semantic Web technologies with machine learning (Breit et al., 2023) reports that only 29 (about 6%) mention using semantic processing modules of some kind (where, by ‘semantic’, the study means symbolic knowledge representation). The dominant use cases for such modules relate to rulesets (learning them, improving and applying them) and to data enrichment. The study also finds that of these 29 papers, only 20 (about 4% of the total) mention using reasoning capabilities to infer knowledge. We consulted that study’s companion SWA and machine learning systems KG resource, SWeMLS-KG (Ekaputra et al., 2023), to find those 4% of papers. Of the 17 we identified, we found only five to use OWL reasoning. Another recent survey and vision paper (d’Amato et al., 2023) reviews the role of KGs in machine learning, pointing out gaps and opportunities, and also observes that KG symbolic reasoning methods are under-explored and largely disregarded.

One factor explaining this under-exploration may be the cross-disciplinary nature of the endeavour. NeSy research with OWL-based KG reasoning requires researchers to be familiar not just with deep learning, KGs, and logic, but with Semantic Web technologies, especially OWL ontologies, as well. The authors of d’Amato et al. (2023) point to the prevalence of huge public KGs and to the perceived scalability limitations of symbolic reasoning methods in the face of such large KGs as an explanation. A third factor may be that the possibility of using OWL-based KG technologies to tailor symbolic deduction engines for NeSy systems has been under-recognised. After all, the Semantic Web was not conceived with such applications in mind.

The objective of this article (which builds on Herron et al., 2023) is to argue the case for NeSy AI research using OWL-based KGs. We hope to inspire more NeSy research using OWL-based KGs by raising awareness of their benefits, capabilities and flexible applications, especially with respect to reasoning. OWL-based KGs are exemplars of the symbolic knowledge representation and symbol manipulation and reasoning machinery that critiques of deep learning, such as by Chollet (2018), Kautz (2022) and Marcus (2018a; 2018b; 2020), advocate be incorporated in NeSy systems. By drawing upon illustrative examples from our own research in visual relationship detection as well as from the literature, and by describing promising research directions, we hope to convince readers of the potential of OWL-based KGs. We also discuss how to enable more NeSy research using OWL-based KGs through the creation of resources such as the recently contributed NeSy4VRD (neurosymbolic AI for visual relationship detection). NeSy4VRD represents one step towards addressing the scarcity of the specialised resources required for NeSy AI research using OWL-based KGs: resources that combine datasets for neural learning with companion OWL ontologies that describe the domain of the data in order to support pertinent symbolic reasoning.

## 2 Benefits and Capabilities of OWL-Based KGs

In this section, we describe benefits and capabilities of OWL-based KGs. We illustrate capabilities by giving examples showing how and why OWL-based KGs can be utilised in NeSy systems.



**Figure 1.** An example neurosymbolic system architecture for detecting visual relationships in images. By using a Web Ontology Language (OWL)-based knowledge graph with an appropriate ontology as a symbolic deduction engine, feedback from OWL reasoning can influence loss to guide neural learning.

## 2.1 Open Standards and Reusable Resources

OWL (the Web Ontology Language) is a key component of the W3C open standards ecosystem of the SW (Berners-Lee et al., 2001; Hitzler et al., 2010; The Semantic Web Stack, 2022; The Semantic Web Wiki, 2001). Open standards facilitate interoperability and promote development of reusable, often free, software resources that make it easy to work with OWL ontologies and OWL-based KGs. Amongst the many such resources are (i) public SW KGs like DBpedia (Lehmann et al., 2015), Wikidata (Vrandečić & Krötzsch, 2014) and Yago (Tanon et al., 2020); (ii) public repositories of curated OWL ontologies like BioPortal (Whetzel et al., 2011) and OBO Foundry (Jackson et al., 2021) in the biomedical domain; (iii) RDF stores like GraphDB (it is not open, but it has a free version) (Ontotext GraphDB, 2023) and RDFox (it is not open, but it has a free academic license) (Nenov et al., 2015); and (iv) efficient OWL reasoners like HermiT (Glimm et al., 2014), Pellet (Sirin et al., 2007), RDFox, and ELK (Kazakov et al., 2011).

## 2.2 Custom Ontologies and Custom KGs

Reusing state-of-the-art ontologies and/or public KGs is a ‘good practice’ option. But researchers can also design their own custom, domain-specific OWL ontologies tailored to their datasets and unique needs. They can then use these to govern and enable reasoning within custom OWL-based KGs. Custom ontologies can also be aligned with publicly available ontologies to enhance interoperability (Euzenat & Shvaiko, 2013).

This custom approach is the one taken in our research into visual relationship detection in images, for which we designed a custom OWL ontology called VRD-World (Herron et al., 2023). This ontology describes the domain of the everyday images of the VRD dataset (Lu et al., 2016), as reflected in the object classes and relationships referred to in the (subject, predicate, object) visual relationships annotated for the images. As depicted in Figure 1, the VRD-World ontology can govern a custom OWL-based KG in the hybrid NeSy systems with which we explore using symbolic reasoning to guide neural learning. While designing the VRD-World ontology, guidance was taken from the large literature on ontology engineering (e.g. Allemang et al., 2020; Keet, 2020; Kendall & McGuinness, 2019; Noy & McGuinness, 2001). The ontology was specified using the free ontology editor Protégé (Musen, 2015), taking advantage of free Protégé plug-in utilities designed to support ontology development, such as ontology debuggers. Many machine learning tools exist to support various different aspects of ontology development such as, for example, concept learners (see d’Amato, 2020).

We designed two versions of a class hierarchy for use with our VRD-World ontology. Both hierarchies contain classes that map to the broad range of everyday object classes present in the images and annotations of the VRD dataset (e.g. person, dog, jacket, surfboard, etc.). The design of one version is entirely custom, and the dataset classes feature exclusively as leaf nodes in the hierarchy. The design of the other version was inspired by Wikidata (Vrandečić & Krötzsch, 2014). We first aligned the dataset classes with matching classes present in the Wikidata class hierarchy. Then, for each such class, we selected a small number of subsumption paths (leading from that class upwards to the top-most class) from the Wikidata hierarchy for inclusion in our class hierarchy. The result is a class hierarchy that is a faithful, tractable subset of the vast Wikidata class hierarchy.

### 2.3 KG Embeddings, KG Completion and Knowledge Injection

KGs (of all kinds) have inspired a large amount of NeSy research into encoding KG symbolic background knowledge into vectors as *KG embeddings*. The embeddings preserve semantic similarity and reflect this similarity by proximity within the embedding vector space (Chen et al., 2020; d’Amato, 2020; Dai et al., 2020; Nickel et al., 2015; Rossi et al., 2021; Wang et al., 2021). The primary application area of KG embeddings so far has been tasks relating to *KG completion: link prediction* (relating individuals in a KG) or *type prediction* (classifying individuals in a KG). Regardless of the model used to generate the embeddings (of which there are many), these link and type prediction problems are typically cast as neural classification problems, where the embedded KG knowledge is used for training and where methods exploiting the proximity principle are applied (as a form of geometric inference) to help make predictions.

Like all KGs, OWL-based KGs are readily used in NeSy research that leverages KG embeddings. OWL2Vec\* (Chen et al., 2021) is one embedding model designed for this purpose. Notice, though, that these applications of KG embeddings focus on leveraging KG symbolic background knowledge only. So even if the KG in question is OWL-based, its reasoning capabilities are generally not employed in these applications.

*Link inference* and *type inference* performed by logical reasoning are, however, the bread and butter of OWL reasoners. When an OWL reasoner infers the knowledge that is entailed by the inference semantics of a governing OWL ontology in the presence of KG data, it completes the KG by introducing new, explicit (knowledge) triples that were previously implicit. This process is called materialisation. The logical soundness of these inferences is guaranteed, whereas embedding-based KG completion is approximate and potentially error prone. The extent to which the KG is extended (completed) is commensurate with the richness of the inference semantics of the governing ontology and the nature of the KG data present at the time of materialisation. Our point is that OWL-based KGs can add important value in any NeSy task associated with KG completion. OWL reasoning can be used to complete a KG automatically, as far as possible, and then NeSy KG completion (NN emulated reasoning) can be used for special cases that the OWL ontology in question does not address or that OWL cannot address in general.

As discussed by Buffelli and Tsamoura (2023), the phrase ‘knowledge injection’ is used with different meanings. It can be used to refer to the injection of knowledge in symbolic form, such as in logic tensor networks (LTNs) (Badreddine et al., 2022; Serafini & d’Avila Garcez, 2016), with its *Real Logic* axioms that are woven into loss functions. It can also refer to the injection of knowledge represented in subsymbolic form, e.g. as embeddings. One subcategory of subsymbolic knowledge injection involves the use of KG embeddings (e.g. Myklebust et al., 2022) as domain knowledge supplements to primary training data. The hypothesis here is that ‘data + knowledge’ can enhance deep learning. This is what the authors of Fu et al. (2023) mean when they speak of knowledge injection. According to Sheth et al. (2019), this approach is called knowledge-infused learning. Much of the research in this area focuses on language models. Using KG entity linking techniques, concepts and entities mentioned in text are matched with corresponding KG entities. The (pre-computed) embeddings of matched KG entities are then looked-up and injected into language models (typically during fine-tuning), often deep within transformer self-attention blocks. For example, (Peters et al., 2019; Roy et al., 2023; Yamada et al., 2020; Zhang et al., 2019) explore variations on this theme and all report performance improvements from injecting KG embeddings as background knowledge.

The efficacy of this approach to KG knowledge injection has come into question, however. Having examined several established language model knowledge injection frameworks and repeated the published experiments, the authors of Fu et al. (2023) conclude that whatever is injected has an effect indistinguishable from that of Gaussian noise. They surmise that fine-tuning does not permit pre-trained language models sufficient opportunity to disentangle and assimilate the latent knowledge in the injected KG embeddings.

We suspect that the prevailing paradigm of matching text to KG entities, and injecting the corresponding embeddings, may limit the use of KG knowledge. KGs express knowledge relationally and KG embedding spaces attempt to represent that relational knowledge subsymbolically. The knowledge is implicit in the relations between pairs and clusters of vectors in the embedding space, and in how clusters are distributed across the space. There may be little usable knowledge residing in individual embeddings considered in isolation (i.e. single points in space). Discordance between heterogeneous text and KG embedding spaces may also be a factor contributing to the findings by Fu et al. (2023). We suggest that new paradigms crafted to facilitate the harvesting and injection of relational knowledge warrant exploration.

Advances in subsymbolic KG knowledge injection will in part be driven by advances in KG embedding models. As discussed in papers such as Abboud et al. (2020), Chen et al. (2023) and d’Amato et al. (2023), KG embedding models currently capture only a portion of the rich semantics of OWL ontologies. Some embedding models focus on capturing lexical and syntactic patterns such as entity/word correlations; others focus on capturing aspects of logical relationships such as hierarchical structure; integrating numeric literals remains a challenge. Research into methods that can embed

more of OWL’s logical expressiveness and logical relationships, and that can embed the multiple aspects mentioned here jointly, remains preliminary (e.g. Jackermeier et al., 2024).

With any application of KG embeddings – KG completion, knowledge injection deep into NNs, or otherwise – the reasoning capabilities of OWL-based KGs may potentially add important value. It seems intuitive that fully materialised OWL-based KGs, where everything implicit has been made explicit, contain more embeddable knowledge and should lead to embedding spaces that better reflect the totalities of symbolic domain knowledge entailed by KGs. However, more embeddable knowledge (more triples) does not necessarily lead to better performance in downstream tasks that consume KG embeddings. Studies such as Iana and Paulheim (2020) show that downstream performance may degrade. Nevertheless, the authors of d’Amato et al. (2023) share this intuition and call for extensive research to study the potential of materialised KGs in KG embeddings, so that guidance can evolve around whether, when, and why to use KG materialisation.

## 2.4 OWL-Based KG Reasoning, Rules, and Symbolic Deduction Engines

Despite recent successes, large language models are notorious for their lack of reliability in reasoning. In contrast, the reliability of OWL reasoning is guaranteed because it is grounded in formal Description Logics (DLs). DLs are decidable fragments of first-order logic with strong connections to set theory (Baader et al., 2007, 2017; Brachman & Levesque, 2004; Nardi & Brachman, 2003). A prominent example is *SROIQ* (Horrocks et al., 2006), the highly expressive DL that underlies that latest version of OWL, OWL 2.

OWL reasoning falls into two broad categories. One category involves the inference of new knowledge, where entailed but implicit triples are made explicit by materialising them (inserting them into the KG). The other category involves checking the logical consistency of ontologies and KGs. Both categories rely on logical inference rules, but whereas the rules of the former category have logical consequents, the rules of the latter category do not. Both categories of reasoning are commonly used for debugging during OWL ontology development (Jiménez-Ruiz, 2010; Kejriwal et al., 2021). What appears to be less well recognised is that both categories of OWL reasoning can also be leveraged in NeSy systems. Some RDF triple stores (like GraphDB and RDFox) do their OWL reasoning as each triple is inserted into the KG. This has important implications for the feasibility of integrating OWL-based KG reasoning into NeSy systems. New knowledge that is entailed by the insertion of a triple (or a small set of triples) is immediately available for inspection, and triple insertion attempts that violate logical consistency rules are immediately rejected (so as to maintain the overall logical consistency of the KG). Both types of response (or feedback) provide information that can be leveraged in the context of hybrid NeSy systems. As is depicted in Figure 1, such OWL-based KG technologies can be used to assemble symbolic deduction engines that perform domain-specific and system-specific OWL reasoning on-call and on-the-fly. Figure 1 shows the feedback from on-the-fly OWL-based KG reasoning being used to guide neural learning by influencing the calculation of loss.

Depending on the application, the KGs of such symbolic deduction engines may need to contain only minimal data at any one moment. As long as a governing OWL ontology is present, fast OWL reasoning can proceed against small number of inserted data triples, and these might be deleted just as rapidly once the momentary reasoning service has been performed and the feedback delivered. OWL reasoning in such symbolic deduction engines can be leveraged not just during NN training but during inference as well. For instance, predictions of visual relationships generated at inference time can be inserted into a KG in order to verify their semantic validity. Ones found to be semantically invalid (i.e. that lead to logical inconsistency) can be filtered from the set of predictions.

Before we discuss examples of NeSy systems found in the literature that use OWL reasoning, we first review some basic OWL reasoning examples. Link inference is driven mainly by the inference semantics associated with the characteristics (e.g. symmetry and transitivity) and relationships (e.g. inverses, subproperties and equivalent properties) declared for the object properties of an OWL ontology. Suppose a property `beside` is declared to be symmetric, and a property `over` is declared to have as inverse the property `under`. In this case, given triples `(a, beside, b)` and `(c, over, d)`, OWL reasoning infers `(b, beside, a)` and `(d, under, c)`. Type inference (subsumption reasoning) is driven mainly by the class hierarchy of an ontology through the transitive property `rdfs:subClassOf`. If the ontology declares `(E, rdfs:subClassOf, F)` and data asserts `(e, rdf:type, E)`, OWL reasoning infers `(e, rdf:type, F)`.

The authors of Chang et al. (2020) leverage OWL type inference in a more elaborate way such that a tutoring system can react intelligently in response to interactions with human learners. A custom OWL ontology models the tutoring system domain and contains descriptions of classes that correspond to tutoring system actions. Data regarding learner interactions with the system are progressively loaded into the system’s KG. The OWL reasoner *HermiT* reasons over the KG to infer new knowledge based on the learner interaction data. In the process, each learner interaction is classified as belonging to one of the tutoring action classes, resulting in inferred triples such as, say, `(learnerX-interactionN, rdf:type,`

GiveEncouragement). The inferred classification triples are interpreted as predictions of the best next action for the tutoring system to take. The mechanism works because OWL allows a class (concept) to be described in terms of the characteristics possessed by the members of the class. This class description capability is a defining feature of DLs.

OWL reasoning can be extended by accompanying OWL ontologies with complementary rules. Such rules are constructed with reference to the classes and properties defined in the ontology, and they typically infer new knowledge (new triples). Various rule technologies for extending OWL exist, such as SPARQL rules (Harris & Seaborne, 2013), SWRL (Semantic Web Rule Language) rules (Horrocks et al., 2004), and (more commonly, today) Datalog rules (Abiteboul et al., 1995; Ceri et al., 1989; Green et al., 2013).

One example in the literature, (Donadello et al., 2019), uses a custom OWL ontology that describes dietary and physical activity domains and healthy lifestyle behaviours, supplemented with SPARQL rules representing unhealthy lifestyle behaviours, in a digital healthcare NeSy system. User diet and activity data are loaded into the system’s KG. The RDFpro tool (Corcoglioniti et al., 2015) drives the reasoning and rule execution. If a SPARQL rule is satisfied, an unhealthy behaviour has been detected, and the rule infers (creates) instances of rule violations in the KG. The rule violations are then rendered into natural language to encourage healthier user behaviours. The authors of Mouakher et al. (2019) use a custom OWL ontology supplemented with SWRL rules as part of a system for monitoring vineyards. Data gathered by a wireless sensor network measuring micro climate conditions around a vineyard are funnelled into the system’s KG. The Pellet OWL reasoner reasons over the KG to infer new triples that are interpreted as predictions of risk of particular diseases and pests. The authors of Nakawala et al. (2019) use a custom OWL ontology supplemented with SWRL rules as part of a NeSy system for recognising surgical processes for robot-assisted surgery. A CNN recognises the current surgical workflow step; an LSTM RNN predicts the next surgical workflow step; and by reasoning over these inputs in the presence of the ontology and SWRL rules, the Pellet OWL reasoner infers supplementary surgical context information, such as the surgical phase, surgical instruments used, and actions to be taken.

The basis of the logical knowledge representation formalism for combining OWL ontologies with Datalog rules is defined by Grosz et al. (2003). This formalism permits rules that refer to ontology vocabulary to be layered on top of ontologies, and it permits logic programming algorithms to reason efficiently over large ontologies. In the VRD dataset setting, a Datalog-like rule describing when it is reasonable to infer the visual relationship  $(x, \text{wear}, y)$  might be represented as follows:

$$\text{wear}(x,y) := \text{WearCapableThing}(x), \text{WearableThing}(y), \text{ir}(y,x) > 0.8$$

In the body of this rule, the first two conditions rely on leveraging OWL type inference; the third condition extends OWL’s capabilities by evaluating the spatial relationship between the bounding boxes of the two objects. Suppose an object detection NN predicts  $(x, \text{rdf:type}, \text{Dog})$  and  $(y, \text{rdf:type}, \text{Jacket})$ . OWL reasoning will infer the membership of these individuals in all of the parent classes of Dog and Jacket and, in so doing, determine whether or not  $(x, \text{rdf:type}, \text{WearCapableThing})$  and  $(y, \text{rdf:type}, \text{WearableThing})$ . If both of these conditions are satisfied, and the bounding box for  $y$  is mostly enclosed within the bounding box for  $x$  (as measured by function  $\text{ir}()$ ), then the body of the Datalog rule is satisfied and the visual relationship in the head of the rule is inferred:  $(x, \text{wear}, y)$ . We intend to explore the effect of such rules using tools such as RDFox, which translates OWL axioms into Datalog rules and performs all of its OWL reasoning using Datalog reasoning algorithms. This allows for seamless blending of reasoning over the OWL 2 RL profile with reasoning over supplementary Datalog rules that extend OWL reasoning.

### 3 Promising Research Directions

In this section, we describe some application areas where the potential for leveraging OWL-based KGs and OWL reasoning in NeSy systems looks promising.

#### 3.1 Using OWL Reasoning to Enhance Annotations and Strengthen Weak Labelling

OWL reasoning has been used to enhance annotations. The authors of Edris et al. (2017) use a small ontology crafted using a fuzzy DL, and a reasoning engine that uses a tableaux algorithm for doing *SHOIQ* DL reasoning (which is virtually OWL reasoning), to semantically enhance the annotations of images. Their knowledge-based framework detects (infers) what they call implicit, ‘semantic context’ concepts. For example, the co-occurrence of ‘sea’ and ‘tree’ in an image might lead to the inference of the context concept ‘beach’, which enhances the semantic interpretation of the image. The authors of Ballan et al. (2010) use a custom OWL ontology (based on a subset of WordNet), supplemented with SWRL rules (learned from data), to enhance video annotations. The SWRL rules infer instances of new, higher-level, composite

concepts relevant to videos. For example, a video containing furniture, lamps, computers, etc., might be inferred to be ‘indoors’.

The VRD dataset we use in our research has weak labelling in the sense that its images are not exhaustively annotated, either in terms of objects or relationships. The visual relationship annotations of its images are sparse and arbitrary, and hence the supervision they provide during NN training is partial and inconsistent across images. OWL reasoning can mitigate weak labelling by augmenting it and making it more consistent. The properties of our VRD-World ontology (which correspond to predicates in annotated visual relationships) are rich in characteristics and relationships that carry inference semantics for link prediction (see Section 2.4). Our experiments with OWL reasoning over VRD-World have shown that the average number of annotated visual relationships per VRD training image increases by a factor of 2.5. Supplementing the ontology with Datalog rules that extend OWL reasoning is expected to yield further augmentation.

The augmentation of the ground-truth annotations for each image could, in theory, be performed in real-time within a NN training loop, but the same reasoning and augmentation would be performed repeatedly, each time each image is re-encountered in successive training epochs. So, in our case, it is more efficient computationally to do the annotation augmentation once, upfront, by materialising a KG containing the annotations of all images and saving the augmented annotations to a file. But the option of using OWL reasoning to perform annotation augmentation on-the-fly, is worthy of note because settings may arise where this facility is advantageous. Either way, augmented, denser and more consistent annotations are likely to provide a less noisy loss signal for neural learning.

The examples just discussed share the notion of using OWL reasoning to infer plausible annotations in the absence of explicit annotations. This notion has broad application. Within supervised learning, it may apply to many datasets (like the VRD dataset) that are not exhaustively annotated. It is also relevant in semi-supervised learning (where some examples are labelled, others not), and potentially in unsupervised learning problems as well. Further, the notion of inferring plausible annotations may be valuable in  $k$ -shot learning scenarios, supervised or otherwise. For example, in the VRD dataset, zero-shot cases arise where objects and relationships that conform to particular visual relationship types, for example, (person, ride, elephant), are present in both training and test images, but where only the test instances have been given ground-truth visual relationship annotations. If, during training, OWL reasoning infers that predictions of such relationships are plausible, then, despite the absence of ground-truth annotations attesting to their validity, we can avoid penalising loss for such predictions and thereby (hopefully) increase the likelihood that the trained NN will predict such relationships when it encounters them in test images.

### 3.2 Enabling NNs to Emulate OWL Reasoning

One approach to NeSy AI involves introducing structural extensions to NN architectures and injecting background knowledge as strong priors in weight matrices. An example of this approach is Kopparti and Weyde (2019). As part of our research, we have begun to explore this approach to NeSy AI by considering the feasibility of transferring OWL-based KG knowledge to NNs (or otherwise equipping them) so that they might emulate aspects of OWL reasoning. One idea involves representing the transitive closure of an OWL ontology class hierarchy (as inferred by OWL reasoning) as a binary adjacency matrix (per graph theory) that can be injected into a classification NN as the weight matrix of a supplementary classification output layer. A proof of concept exercise indicates that this tactic should permit a classifier to emulate the subsumption reasoning (type inference) capabilities of an OWL-based KG. For example, suppose a NN classifier that initially classifies a data sample  $x$  as being a Dog. The structural extension (i.e. the additional classification layer) could then generalise (or pseudo-infer) that  $x$  is thus also a Carnivore, a Mammal, an Animal, a LivingThing, etc., just as would an OWL reasoner in an OWL-based KG governed by the same ontology, were the triple ( $x$ , `rdf:type`, Dog) inserted.

The potential for leveraging cross-over synergies such as this between OWL-based KGs and NNs is ripe for exploration and development. One avenue for investigation might involve adding more learnable layers to a classifier following the class generalisation extension layer just described so that learning might proceed driven by generalised class predictions. Another option might be to extend the solution just described for generalising multi-class, single-label predictions to the multi-class, multi-label setting. A further option is to apply this technique for transferring OWL class hierarchy knowledge to a NN to the transference of OWL property hierarchy knowledge to a NN. Our Predicate Prediction NN (in Figure 1), for example, predicts predicates (which map to OWL properties) that relate two objects. The VRD-World ontology declares that property `sitUnder` is a subproperty of `under`, and that `under` is a subproperty of `below`. If a NN prediction ( $x$ , `sitUnder`,  $y$ ) is inserted into an OWL-based KG governed by VRD-World, OWL reasoning will infer triples (visual relationships) ( $x$ , `under`,  $y$ ) and ( $x$ , `below`,  $y$ ). But an adjacency matrix encoding the transitive closure of the VRD-World object property hierarchy should enable a NN to make these two generalised visual relationship predictions for itself.

### 3.3 Using OWL Reasoning for Applying Logical Constraints

Much NeSy research explores using background knowledge expressed in first-order logic, propositional logic or logic programming as constraints to guide neural learning, often by manipulating loss to encourage constraint satisfaction. Examples are (i) the NN training framework LTN which allows fuzzy, first-order *Real Logic* knowledge axioms (constraints) to be defined over training data; (ii) the set of propositional logic constraints specified for the ROAD-R dataset (Giunchiglia et al., 2023); and (iii) the (Prolog) rules defined by Cornelio et al. (2023). OWL-based KG technologies can be used for the same purpose. Recall from Section 2.4 that OWL reasoning spans two categories of logical inference, one which infers new knowledge and one which checks for logical consistency. Both categories of inference can be leveraged to enable OWL-based KGs to participate in research associated with the *logical constraints* approach to NeSy AI.

First we consider the latter category. We assume a context of on-the-fly OWL reasoning, which permits logical consistency checks to be applied at the point of triple insertion. In this setting, OWL’s logical consistency inference rules can be leveraged as though they were logical constraints. If an OWL-based KG rejects the insertion of a triple, this event signals a violation of a consistency rule and therefore the violation of some logical constraint. Such events can thus be used to penalise loss. The number and nature of the logical constraints covered by OWL’s logical consistency rules varies with the design of the OWL ontology. If one opts to design a custom ontology, one can arrange for specific logical constraints to exist that will be covered automatically by OWL’s logical consistency checks, much like one might craft a logic programme with specific inference intentions in mind. For a given rejection event, it may be feasible to surmise which logical constraint was violated based upon knowledge of the ontology’s design and of the triple(s) that triggered the violation. Alternatively, the OWL reasoner Pellet provides explanations for logical inconsistencies to an extent. The Protégé editor also has facilities and plug-ins that provide explanations for detected ontology inconsistencies.

As a concrete example, we discuss the use of domain and range restrictions as logical constraints in connection with the VRD dataset. We describe how these have been used in the context of LTN and then compare that approach with how they can be used in OWL ontologies. According to Donadello and Serafini (2019), negative domain and negative range LTN *Real Logic* axioms (constraints) are used to train binary classifiers for the predicates of the VRD dataset. The VRD dataset has 100 object classes, so to train a binary classifier for predicate *wear*, say, close to 100 negative domain LTN constraints, such as

$$\forall xy \text{ wear}(x, y) \rightarrow \neg \text{Laptop}(x) \quad \forall xy \text{ wear}(x, y) \rightarrow \neg \text{Sofa}(x) \quad \forall xy \text{ wear}(x, y) \rightarrow \neg \text{Tree}(x) \dots$$

would have been required to enumerate the objects that cannot be in the domain of predicate *wear*. Similarly, close to 100 negative range LTN constraints, such as

$$\forall xy \text{ wear}(x, y) \rightarrow \neg \text{Table}(y) \quad \forall xy \text{ wear}(x, y) \rightarrow \neg \text{Car}(y) \quad \forall xy \text{ wear}(x, y) \rightarrow \neg \text{Oven}(y) \dots$$

would have been required to enumerate what cannot be in the range.

OWL can express equivalent logical constraints, and can do so more concisely. The VRD-World ontology can express the equivalent of the close to 100 negative domain constraints (1) by defining the (disjoint) classes `WearCapableThing` and `WearIncapableThing` in its class hierarchy, and (2) by declaring that the domain of object property *wear* is restricted to members of the class `WearCapableThing`, with the OWL axiom

```
vrd:wear rdfs:domain vrd:WearCapableThing.
```

Similarly, the logical constraint equivalent of the close to 100 negative range LTN *Real Logic* axioms can be expressed (1) by defining the (disjoint) classes `WearableThing` and `NonWearableThing` in the class hierarchy of VRD-World, and (2) by declaring that the range of object property *wear* is restricted to members of the class `WearableThing`, with the OWL axiom

```
vrd:wear rdfs:range vrd:WearableThing.
```

Figure 1 shows how an OWL-based KG with an appropriate ontology (such as VRD-World) can be used, in the guise of a symbolic deduction engine, to leverage ontological rules as logical constraints to guide neural learning. Suppose the Object Detection NN predicts that *x* is a dog and *y* is a surfboard. If the multi-class, multi-label Predicate Prediction NN shows a tendency to predict a visual relationship such as (*dog*, *wear*, *surfboard*), the RDF triples representing this prediction

```

vrd:x rdf:type vrd:Dog
vrd:y rdf:type vrd:Surfboard
vrd:x vrd:wear vrd:y

```

can be inserted into the KG. OWL type inference will infer that  $x$  is a `WearCapableThing` (i.e. in VRD-World, dogs can wear things), and that  $y$  (a surfboard) is a `NonWearableThing`. But the axiom expressing the range restriction on predicate `wear` will lead OWL to infer that  $y$  is a `WearableThing`. OWL’s logical consistency checks will detect that individual  $y$  is a member of both `WearableThing` and `NonWearableThing`, two classes declared to be disjoint which cannot share members. OWL reasoning will detect and report this logical inconsistency (this constraint violation). This feedback can be used to penalise loss to help the Predicate Prediction NN learn to avoid predicting visual relationships that are semantically invalid.

In addition to illustrating that OWL-based KGs can emulate the logical constraints approach to NeSy AI, this example also illustrates an important advantage possessed by OWL-based KGs over other approaches to using logical constraints. The research by Donadello and Serafini (2019) shows that the logical constraints approach to NeSy AI is exposed to the risk of combinatorial explosion, where the number of constraints requiring expression grows too rapidly with the number of classes in the dataset. Almost 200 LTN *Real Logic* axioms would have been needed in relation to just one VRD predicate, `wear`. And about 30 of the 70 VRD predicates admit domain and/or range restrictions of some kind. Indeed, (Donadello & Serafini, 2019) reports implementing a ‘tractable sample’ only of the LTN *Real Logic* axioms implied by the negative domain/range constraints training strategy selected for the experiments. In contrast, once an appropriate class hierarchy is defined, expressing powerful domain and range restrictions in OWL is easy.

This comparative advantage possessed by OWL for expressing background knowledge (and logical constraints) concisely is reinforced by considering a different example. The autonomous vehicle driving videos and annotated bounding boxes of the ROAD-R dataset (Giunchiglia et al., 2023) are accompanied by 243 manually specified propositional logic constraints that define the permissible combinations of labels for 10 agent classes, 19 agent action classes and 12 agent location classes. Amongst the 243 logic constraints, 45 have a format such as  $\neg\text{Car} \vee \neg\text{Bus}$  (meaning ‘an agent cannot be a car and a bus at the same time’), or  $\neg\text{RedTL} \vee \neg\text{GreenTL}$  (meaning ‘a traffic light cannot be red and green at the same time’). Collectively, these 45 constraints express mutual exclusiveness between pairwise combinations of the 10 agent classes. Precise counterparts of these 45 propositional constraints can be represented in OWL with just two axioms, such as

```

:DisjointAgents rdf:type owl:AllDisjointClasses.
:DisjointAgents owl:members (:Car :Bus :Motorbike :Pedestrian...).

```

Similarly, 66 of the ROAD-R propositional constraints express pairwise mutual exclusiveness amongst the 12 agent location classes. Counterparts of these can be represented in OWL using just two more such axioms.

In fact, by making appropriate use of OWL’s constructs for declaring domain and range restrictions, disjoint classes, disjoint properties, functional properties, and the like, it may well be possible to design an OWL ontology that emulates all of the 243 propositional logic constraints specified for the ROAD-R dataset. Doing so would, in theory, make it feasible to repeat the ROAD-R experiments by Giunchiglia et al. (2023) using an OWL-based KG as a symbolic deduction engine instead of using the original propositional constraints with a SAT solver as a reasoning engine.

Now we consider how OWL’s other category of logical inference – the one that infers new knowledge – can be leveraged in the context of using logical constraints to guide neural learning. An alternate strategy for using OWL to emulate the propositional logical constraints of the ROAD-R dataset is to employ the concept of *integrity constraints* described by Kharlamov et al. (2016). Rather than checking and enforcing ontology (KG) logical consistency, integrity constraints employ Datalog rules that supplement an OWL ontology to represent logical constraints. Such constraint rules, if satisfied, infer explicit new triples into the KG which can then be queried and interpreted as signals of constraint violations. For example, the Datalog integrity constraint rule

```

Violation(X) :- Car(X), Bus(X)

```

declares that an agent cannot be both a car and a bus. Similarly, the hypothetical ontology would permit the creation of rules like

```

Violation(X) :- TL(X), hasL(X, Y), hasL(X, Z), GreenL(Y), RedL(Z)

```

to establish that a traffic light (TL) cannot be red and green at the same time. Note that, unlike in the propositional case, Datalog rules provide additional granularity for describing cases in which an image contains more than one TL. This approach using integrity constraints could also be applied to the VRD case, with rules like

```
Violation(X) :- vrd:wear(X, Y), not vrd:WearCapableThing(X)
Violation(Y) :- vrd:wear(X, Y), not vrd:WearableThing(Y)
```

### 3.4 Integrating OWL-Based KG Reasoning With Existing NeSy Frameworks

OWL-based KG symbolic knowledge and deductive reasoning can be integrated with and leveraged by existing logic-based NeSy frameworks such as LTN. LTN functions that encapsulate interactions with OWL-based KGs can, in theory, participate in LTN *Real Logic* knowledge axioms used to train NNs. One precondition is that there is sufficient contextual information contained in LTN tensors (or otherwise) to permit RDF triples to be constructed and inserted into a KG to drive reasoning, and/or to enable KG queries to be formulated. The only other precondition is that the results of KG interactions can be mapped to fuzzy truth values in  $[0, 1]$ .

One application of this idea involves using OWL-based KG reasoning to manage the risk of combinatorial explosion (described in Section 3.3) to which the *logical constraints* approach to NeSy AI is exposed. A prime cause of exposure to this risk derives from the fact that logical constraints (as used by LTN and the ROAD-R dataset, for example) are restricted to being expressed in terms of the low-level, granular object classes present in data and their annotations. The option to express constraints more concisely, in terms of higher-level, more general classes, is not available. In contrast, OWL ontologies routinely possess rich class hierarchies that permit ontological rules to be defined in terms of high-level, general classes, which affords simplicity and parsimony.

To illustrate, consider again the research undertaken by Donadello and Serafini (2019), where the use of negative domain and range constraints leads to the need for an intractable number of LTN *Real Logic* knowledge axioms to be crafted. This time, however, suppose that we integrate interactions with an OWL-based KG (acting as a symbolic deduction engine) into our LTN *Real Logic* knowledge axioms in order to map the granular classes present in the data to higher-level, more general classes defined in the class hierarchy of the VRD-World ontology. Using this strategy, we can imagine replacing the original (close to) 200 *negative* domain and range LTN constraints used to train a binary classifier for VRD predicate *wear* with just two *positive* LTN *Real Logic* knowledge axiom constraints, such as

$$\forall xy \text{ wear}(x, y) \rightarrow \text{WearCapableThing}(x) \quad \forall xy \text{ wear}(x, y) \rightarrow \text{WearableThing}(y)$$

For clarity, note that, unlike in OWL reasoning and Datalog rule reasoning where new knowledge is inferred, LTN *Real Logic* knowledge axioms like these do not infer anything, despite the fact that they express logical implications. Instead, the degree to which LTN *Real Logic* knowledge axioms are satisfied is measured, and the extent to which they are not satisfied represents loss that drives neural learning.

A more compute-efficient implementation of the proposal just described is also feasible. In this setting, we only wish to exploit the type inference (subsumption reasoning) capabilities of OWL reasoning. But, as we saw in Section 3.2, for a given OWL ontology, these capabilities can be fully encoded in the *adjacency matrix* of the transitive closure of the ontology’s class hierarchy. So instead of interacting with an OWL-based KG to use OWL reasoning to map granular classes to higher-level classes, we can instead use the *adjacency matrix* to do the mapping.

Integrating OWL-based KG reasoning with LTN in the manner just described is a specialised approach to blending OWL (a DL) with (fuzzy) first-order logic (FOL). Another approach to blending OWL with FOL is to translate it into FOL. This leads to opportunities to extend OWL reasoning capabilities by augmenting OWL ontologies with supplementary FOL axioms (or ontologies) that express things OWL cannot. Tools that support this approach include *Hets* (The Heterogeneous Tool Set) (Mossakowski et al., 2007) and *Gavel-OWL* (Flügel et al., 2022). The resulting integrated FOL ontologies (OWL-to-FOL axioms, plus supplementary FOL axioms) that such tools produce are reasoned over using established FOL Automated Theorem Provers (ATPs).

The ‘translate OWL to FOL’ strategy just described and the ‘translate OWL to Datalog’ strategy mentioned in Section 2.4 are instances of the same pattern: (i) translate OWL into logic space X; (ii) optionally extend OWL with supplementary knowledge expressible in logic space X; and (iii) reason using the logical inference technology established for logic space X. Such strategies widen the window of opportunity for leveraging the available OWL resources in NeSy frameworks.

## 4 Enabling NeSy Research Using OWL-Based KGs With NeSy4VRD

Sections 2 and 3 focus on inspiring more NeSy research using OWL-based KGs by highlighting their benefits, capabilities, and applications, especially with respect to reasoning, and particularly in symbolic deduction engine settings. But inspiration alone may not be enough. To undertake NeSy research with OWL-based KGs and reasoning in a practical way, researchers need to also be enabled with appropriate dataset resources. Resources are needed that combine data for neural learning with strongly-aligned, companion OWL ontologies describing the domains of the data in order to support directly pertinent symbolic OWL reasoning. Such resources are scarce. We suspect this scarcity represents a silent barrier that inhibits NeSy research using OWL-based KGs that might otherwise be undertaken. As well as echoing our observations, (d’Amato et al., 2023) calls for a central repository for such specialised resources in order to simplify their discovery. One resource of this kind (one which belongs in such a repository) is *NeSy4VRD* (neurosymbolic AI for visual relationship detection). NeSy4VRD was co-developed and published by the authors of this article (Herron et al., 2023) to help address the scarcity issue.

NeSy4VRD consists of the following components and services:

- (i) the images of the original VRD dataset (Lu et al., 2016) (distributed with permission from one of the principals associated with its creation) in order to make them publicly available once again;
- (ii) quality-improved versions of the original VRD visual relationship annotations that have been comprehensively customised and extended to enable the engineering of a robust ontology;
- (iii) a strongly-aligned, custom-designed companion OWL ontology, called VRD-World, that precisely describes the domain of the images and visual relationships;
- (iv) sample Python code for loading the annotated visual relationships into a KG hosting the VRD-World ontology, and for extracting them from a KG and restoring them to their native format;
- (v) support for extensibility of the annotations (and, thereby, the ontology) in the form of (a) comprehensive Python code enabling deep but easy analysis of the images and their annotations, (b) a custom, text-based protocol for specifying annotation customisation instructions declaratively and (c) a configurable, managed Python workflow for customising annotations in an automated, repeatable process;
- (vi) comprehensive documentation describing (a) how to use the extensibility support infrastructure, (b) how to share annotation/ontology extensibility projects undertaken by researchers in pursuit of their private research interests, (c) how to reuse shared extensibility projects and use the NeSy4VRD workflow to compose them in novel combinations and (d) how the ability to undertake, share, reuse and compose NeSy4VRD extensibility projects represents a new model of collaborative data annotation that we call Distributed Annotation Enhancement.

The NeSy4VRD dataset package (VRD images, quality-improved visual relationship annotations and companion VRD-World OWL ontology) is distributed on Zenodo.<sup>1</sup> The NeSy4VRD extensibility support infrastructure and comprehensive documentation are available on GitHub.<sup>2</sup>

## 5 Conclusion


A central concern of NeSy AI research is to explore ways of combining neural learning with symbolic background knowledge and reasoning. OWL-based KGs are exemplars of symbolic knowledge representation and reasoning technology and machinery. They can do everything that general KGs can do in terms of representing symbolic knowledge and generating embeddings, plus they can perform sound deductive reasoning to both infer new knowledge and enforce logical consistency, and they can do so in the guise of symbolic deduction engines. Given these attractive features, OWL-based KGs warrant more research attention from the NeSy community than they have received. Their potential for contributing to NeSy AI is not being fully explored. By describing and illustrating their benefits, capabilities, and flexible applications, we have endeavoured to inspire more such research. By having contributed NeSy4VRD – a specialised and scarce dataset resource – to the NeSy community, we hope to have lowered barriers to entry and thereby enabled more such research. A recent overview of NeSy systems (Sheth et al., 2023) reports success using an OWL-based KG to boost expert user satisfaction with large language model performance. Like us, the authors strongly advocate the use of KGs (general and OWL-based) as symbolic components in NeSy systems.


## Acknowledgements


This work has been partially supported by the Academy of Medical Sciences Network Grant (Neurosymbolic AI for Medicine, NGR1\1857) and the project ‘XAI4SOC: Explainable Artificial Intelligence for Healthy Aging and Social Wellbeing’ funded by the

Agencia Estatal de Investigación (AEI), the Spanish Ministry of Science, Innovation and Universities and the European Social Funds (PID2021-123152OB-C22).

## ORCID iDs

David Herron  <https://orcid.org/0009-0008-2736-6789>

Ernesto Jiménez-Ruiz  <https://orcid.org/0000-0002-9083-4599>

Tillman Weyde  <https://orcid.org/0000-0001-8028-9905>

## Statements and Declarations

### Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Notes

1. <https://doi.org/10.5281/zenodo.7916355>

2. <https://github.com/djherron/NeSy4VRD>

## References

- Abboud, R., Ceylan, I., Lukasiewicz, T., & Salvatori, T. (2020). BoxE: A box embedding model for knowledge base completion. In *Advances in neural information processing systems* (Vol. 33, pp. 9649–9661). Curran Associates, Inc. [https://proceedings.neurips.cc/paper\\_files/paper/2020/file/6dbbe6abe5f14af882ff977fc3f35501-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2020/file/6dbbe6abe5f14af882ff977fc3f35501-Paper.pdf)
- Abiteboul, S., Hull, R., & Vianu, V. (1995). *Foundations of databases*. Addison-Wesley. <http://webdam.inria.fr/Alice/>
- Allemang, D., Hendler, J., & Gandon, F. (2020). *Semantic web for the working ontologist* (3rd ed.). ACM Books.
- Baader, F., Calvanese, D., McGuinness, D., Nardi, D., & Patel-Schneider, P. F. (Eds.). (2007). *The description logic handbook* (2nd ed.). Cambridge University Press.
- Baader, F., Horrocks, I., Lutz, C., & Sattler, U. (2017). *An introduction to description logic*. Cambridge University Press.
- Badreddine, S., d'Avila Garcez, A., Serafini, L., & Spranger, M. (2022). Logic tensor networks. *Artificial Intelligence*, 303. <https://doi.org/10.1016/j.artint.2021.103649>
- Ballan, L., Bertini, M., Bimbo, A. D., & Serra, G. (2010). Video annotation and retrieval using ontologies and rule learning. *IEEE MultiMedia*, 17(4), 80–88. <https://doi.org/10.1109/MMUL.2010.4>
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, 284(5), 34–43. <https://www.jstor.org/stable/pdf/26059207.pdf>
- Bianchi, F., & Hitzler, P. (2019). On the capabilities of logic tensor networks for deductive reasoning. In *Proceedings of the AAAI-MAKE* (Vol. 2350).
- Brachman, R., & Levesque, H. (2004). *Knowledge representation and reasoning*. Morgan Kaufman.
- Breit, A., Waltersdorfer, L., Ekapatra, F. J., Sabou, M., Ekelhart, A., Iana, A., Paulheim, H., Portisch, J., Revenko, A., Teije, A. t., & van Harmelen, F. (2023). Combining machine learning and semantic web: A systematic mapping study. *ACM Computing Surveys*. <https://doi.org/10.1145/3586163>
- Buffelli, D., & Tsamoura, E. (2023). Scalable theory-driven regularization of scene graph generation models. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 37, pp. 6850–6859). <https://ojs.aaai.org/index.php/AAAI/article/view/25839>
- Ceri, S., Gottlob, G., & Tanca, L. (1989). What You always wanted to know about datalog (and never dared to ask). *IEEE Transactions on Knowledge and Data Engineering*, 1(1), 146–166. <https://doi.org/10.1109/69.43410>
- Chang, M., D'Aniello, G., Gaeta, M., Orciuoli, F., Sampson, D. G., & Simonelli, C. (2020). Building ontology-driven tutoring models for intelligent tutoring systems using data mining. *IEEE Access*, 8, 48151–48162. <https://doi.org/10.1109/ACCESS.2020.2979281>
- Chen, Z., Chen, J., et al. (2021). Zero-shot visual question answering using knowledge graph. *CoRR*. <https://arxiv.org/abs/2107.05348>
- Chen, J., Geng, Y., et al. (2021). Low-resource learning with knowledge graphs: A comprehensive survey. *CoRR*. <https://arxiv.org/abs/2112.10006>
- Chen, J., He, Y., Geng, Y., Jiménez-Ruiz, E., Dong, H., & Horrocks, I. (2023). Contextual semantic embeddings for ontology subsumption prediction. *World Wide Web (WWW)*, 26(5), 2569–2591. <https://doi.org/10.1007/s11280-023-01169-9>
- Chen, J., Hu, P., Jiménez-Ruiz, E., Holter, O. M., Antonyrajah, D., & Horrocks, I. (2021). OWL2Vec\*: Embedding of OWL ontologies. *Machine Learning*, 110(7), 1813–1845. <https://doi.org/10.1007/s10994-021-05997-6>
- Chen, Z., Wang, Y., Zhao, B., Cheng, J., Zhao, X., & Duan, Z. (2020). Knowledge graph completion: A review. *IEEE Access*, 8, 192435. <https://doi.org/10.1109/ACCESS.2020.3030076>

- Chollet, F. (2018). Deep learning: Current limits and what lies beyond them. <https://raais.co/speakers-2018>
- Corcoglioniti, F., Rospocher, M., Mostarda, M., & Amadori, M. (2015). Processing billions of RDF triples on a single machine using streaming and sorting. In *Proceedings of the 30th annual ACM symposium on applied computing* (pp. 368–375). ACM. <https://doi.org/10.1145/2695664.2695720>
- Cornelio, C., Stuehmer, J., Hu, S. X., & Hospedales, T. (2023). Learning where and when to reason in neuro-symbolic inference. In *International conference on learning representations*. <https://openreview.net/forum?id=en9V5F8PR->
- Dai, Y., Wang, S., Xiong, N. N., & Guo, W. (2020). A survey on knowledge graph embedding: Approaches, applications and benchmarks. *Electronics*, 9(5). <https://doi.org/10.3390/electronics9050750>. <https://www.mdpi.com/2079-9292/9/5/750>
- d'Amato, C. (2020). Machine learning for the semantic web: Lessons Learnt and Next Research Directions. *Semantic Web*, 11(1), 195–203. <https://doi.org/10.3233/SW-200388>
- d'Amato, C., Mahon, L., Monnin, P., & Stamou, G. (2023). Machine learning and knowledge graphs: Existing gaps and future research challenges. *Transactions on Graph Data and Knowledge (TGDK)*, 1(1), 8:1–8:35. <https://doi.org/10.4230/TGDK.1.1.8>
- Donadello, I., Dragoni, M., & Eccher, C. (2019). Persuasive explanation of reasoning inferences on dietary data. In *Proceedings of the 1st workshop on semantic explainability, co-located with the 18th international semantic web conference (ISWC 2019)* (Vol. 2465, pp. 46–61). CEUR Workshop Proceedings, CEUR-WS.org. [https://ceur-ws.org/Vol-2465/semex\\_paper2.pdf](https://ceur-ws.org/Vol-2465/semex_paper2.pdf)
- Donadello, I., & Serafini, L. (2019). Compensating supervision incompleteness with prior knowledge in semantic image interpretation. In *IJCNN Hungary, July 14–19* (pp. 1–8). IEEE. <https://doi.org/10.1109/IJCNN.2019.8852413>
- Ebrahimi, M., Eberhart, A., Bianchi, F., & Hitzler, P. (2021a). Towards bridging the neuro-symbolic gap: Deep deductive reasoners. *Applied Intelligence*, 51(9), 6326–6348.
- Ebrahimi, M., Sarker, M. K., Bianchi, F., et al. (2021b). Neuro-symbolic deductive reasoning for cross-knowledge graph entailment. In *Proceedings of the AAAI-MAKE* (Vol. 2846). <http://ceur-ws.org/Vol-2846/paper8.pdf>
- Edris, S. S., Zarka, M., Ouarda, W., & Alimi, A. M. (2017). A fuzzy ontology driven context classification system using large-scale image recognition based on deep CNN. In *Sudan conference on computer science and Information Technology (SCCSIT)* (pp. 1–9). IEEE.
- Ekaputra, F. J., Llugiqi, M., Sabou, M., Ekelhart, A., Paulheim, H., Breit, A., Revenko, A., Waltersdorfer, L., Farfar, K. E., & Auer, S. (2023). Describing and organizing semantic web and machine learning systems in the SWeMLS-KG. In *Proceedings of the extended semantic web conference (ESWC)* (Vol. 13870, pp. 372–389). Springer. [https://doi.org/10.1007/978-3-031-33455-9\\_22](https://doi.org/10.1007/978-3-031-33455-9_22)
- Euzenat, J., & Shvaiko, P. (2013). *Ontology matching* (2nd ed.). Springer. ISBN 978-3-642-38720-3.
- Flügel, S., Glauer, M., Neuhaus, F., & Hastings, J. (2022). When One logic is not enough: Integrating first-order annotations in OWL ontologies. *Semantic Web*. <https://www.semantic-web-journal.net/system/files/swj3440.pdf>
- Fu, P., Zhang, Y., Wang, H., Qiu, W., & Zhao, J. (2023). Revisiting the knowledge injection frameworks. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 10983–10997). Association for Computational Linguistics. <https://aclanthology.org/2023.emnlp-main.677>
- Geng, Y., Chen, J., Ye, Z., Yuan, Z., Zhang, W., & Chen, H. (2021). Explainable zero-shot learning Via attentive graph convolutional network and knowledge graphs. *Semantic Web*, 12(5), 741–765. <https://doi.org/10.3233/SW-210435>
- Giunchiglia, E., Stoian, M. C., Khan, S., Cuzzolin, F., & Lukasiewicz, T. (2023). ROAD-R: The autonomous driving dataset with logical requirements. *Machine Learning*. <https://doi.org/10.1007/s10994-023-06322-z>
- Glimm, B., Horrocks, I., Motik, B., Stoilos, G., & Wang, Z. (2014). Hermit: An OWL 2 Reasoner. *Journal of Automated Reasoning*, 53(3), 245–269. <https://doi.org/10.1007/s10817-014-9305-1>
- Green, T. J., Huang, S. S., Loo, B. T., & Zhou, W. (2013). Datalog and recursive query processing. *Foundations and Trends in Databases*, 5(2), 105–195. <https://doi.org/10.1561/19000000017>
- Grosf, B. N., Horrocks, I., Volz, R., & Decker, S. (2003). Description logic programs: Combining logic programs with description logic. In *Proceedings of the twelfth international world wide web conference* (pp. 48–57). ACM. <https://doi.org/10.1145/775152.775160>
- Harris, S., & Seaborne, A. (2013). *SPARQL 1.1 query language*. W3C Recommendation, W3C. <https://www.w3.org/TR/sparql11-query/>
- Herron, D., Jiménez-Ruiz, E., Tarroni, G., & Weyde, T. (2023). NeSy4VRD: A multifaceted resource for neurosymbolic AI research using knowledge graphs in visual relationship detection. *CoRR*. <http://arxiv.org/abs/2305.13258>
- Herron, D., Jiménez-Ruiz, E., & Weyde, T. (2023). On the benefits of OWL-based knowledge graphs for neural-symbolic systems. In *NeSy 2023: 17th international workshop on neural-symbolic learning and reasoning, July 3–5, La Certosa di Pontignano, Siena, Italy* (Vol. 3432). CEUR Workshop Proceedings, CEUR-WS.org. <https://ceur-ws.org/Vol-3432/paper28.pdf>
- Hitzler, P. (2021). Knowledge graphs in neuro-symbolic AI, (Presentation slides). <https://people.cs.ksu.edu/~textildelow/hitzler/pub2/2021-10-AKBC.pdf>.
- Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P. F., & Rudolph, S. (2012). *OWL 2 web ontology language primer*. W3C Recommendation, World Wide Web Consortium. <http://www.w3.org/TR/owl2-primer/>

- Hitzler, P., Krötzsch, M., & Rudolph, S. (2010). *Foundations of semantic web technologies* (3rd ed.). CRC Press.
- Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., de Melo, G., Gutiérrez, C., Kirrane, S., Labra Gayo, J. E., Navigli, R., Neumaier, S., Ngonga Ngomo, A.-C., Polleres, A., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J. F., Staab, S., & Zimmermann, A. (2021). *Knowledge graphs*. Springer. Synthesis Lectures on Data, Semantics, and Knowledge, <https://kgbook.org/>
- Horrocks, I., Kutz, O., & Sattler, U. (2006). The even more irresistible SROIQ. In P. Doherty, J. Mylopoulos, & C. A. Welty (Eds.), *Proceedings, tenth international conference on principles of knowledge representation and reasoning* (pp. 57–67). AAAI Press. <http://www.aaai.org/Library/KR/2006/kr06-009.php>
- Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosz, B., & Dean, M. (2004). *SWRL: A semantic web rule language combining OWL and RuleML*. W3C Member Submission, W3C. <https://www.w3.org/submissions/SWRL/>
- Iana, A., & Paulheim, H. (2020). More is not always better: The negative impact of A-box materialization on RDF2vec knowledge graph embeddings. In *Proceedings of the CIKM workshops* (Vol. 2699). CEUR Workshop Proceedings, CEUR-WS.org. <https://ceur-ws.org/Vol-2699/paper05.pdf>
- Jackermeier, M., Chen, J., & Horrocks, I. (2024). Dual box embeddings for the description logic  $\mathcal{EL}++$ . In *WWW*.
- Jackson, R., Matentzoglou, N., Overton, J. A., Vita, R., Balhoff, J. P., Buttigieg, P. L., Carbon, S., Courtot, M., Diehl, A. D., Dooley, D. M., Duncan, W. D., Harris, N. L., Haendel, M. A., Lewis, S. E., Natale, D. A., Osumi-Sutherland, D., Ruttenberg, A., Schriml, L. M., Smith, B., Stoeckert Jr, C. J., Vasilevsky, N. A., Walls, R. L., Zheng, J., Mungall, C. J., & Peters, B. (2021). OBO foundry in 2021: Operationalizing open data principles to evaluate ontologies. *Database*, 2021. <https://doi.org/10.1093/database/baab069>. <http://obofoundry.org/>
- Jiménez-Ruiz, E. (2010). *Logic-based support for ontology development in open environments* [PhD thesis]. Jaume I University, Spain. <http://hdl.handle.net/10803/10493>
- Kautz, H. (2022). The third AI summer: AAAI Robert S. Engelmore memorial lecture (Vol. 43, pp. 105–125).
- Kazakov, Y., Krötzsch, M., & Simančík, F. (2011). Unchain my *EL* reasoner. In *Proceedings of the 24th international workshop on description logics (DL'11)* (Vol. 745). CEUR Workshop Proceedings, CEUR-WS.org.
- Keet, C. M. (2020). *An introduction to ontology engineering* ('v1.5' ed.). College Publications. <https://people.cs.uct.ac.za/~mkeet/OEbook/>
- Kejriwal, M., Knoblock, C. A., & Szekely, P. (2021). *Knowledge graphs: Fundamentals, techniques, and applications*. The MIT Press.
- Kendall, E. F., & McGuinness, D. L. (2019). *Ontology engineering*. Morgan & Claypool Publishers. Synthesis Lectures on The Semantic Web: Theory and Technology.
- Kharlamov, E., Grau, B. C., Jiménez-Ruiz, E., Lamparter, S., Mehdi, G., Ringsquandl, M., Nenov, Y., Grimm, S., Roshchin, M., & Horrocks, I. (2016). Capturing industrial information models with ontologies and constraints. In P. Groth, E. Simperl, A.J.G. Gray, M. Sabou, M. Krötzsch, F. Lécué, F. Flöck, & Y. Gil (Eds.), *The semantic web - ISWC 2016 - 15th international semantic web conference, Kobe, Japan, October 17-21, 2016, Proceedings, Part II* (Vol. 9982, pp. 325–343). Lecture Notes in Computer Science. [https://doi.org/10.1007/978-3-319-46547-0\\_30](https://doi.org/10.1007/978-3-319-46547-0_30)
- Kopparti, R., & Weyde, T. (2019). Weight priors for learning identity relations. In *Workshop knowledge representation & reasoning meets machine learning at NeurIPS* (pp. 8–14).
- Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., & Bizer, C. (2015). DBpedia - A large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web*, 6(2), 167–195. <http://dx.doi.org/10.3233/SW-140134>
- Lu, C., Krishna, R., Bernstein, M., & Fei-Fei, L. (2016). Visual relationship detection with language priors. In *European conference on computer vision* (pp. 852–869). Springer. <https://cs.stanford.edu/people/ranjaykrishna/vrd/>
- Marcus, G. (2018a). Deep learning: A critical appraisal. *CoRR*. <http://arxiv.org/abs/1801.00631>
- Marcus, G. (2018b). Innateness, AlphaZero, and artificial intelligence. *CoRR*. <http://arxiv.org/abs/1801.05667>
- Marcus, G. (2020). The next decade in AI: Four steps towards robust artificial intelligence. *CoRR*. <https://arxiv.org/abs/2002.06177>
- Mossakowski, T., Maeder, C., & Lüttich, K. (2007). The heterogeneous tool set, hets. In *International conference on tools and algorithms for the construction and analysis of systems* (Vol. 4424, pp. 519–522). Springer. <https://doi.org/10.1007/978-3-540-71209-1>
- Motik, B., Patel-Schneider, P. F., Parsia, B., Bock, C., Fokoue, A., Haase, P., Hoekstra, R., Horrocks, I., Ruttenberg, A., Sattler, U., & Smith, M. (2012). *OWL 2 web ontology language: Structural specification and functional-style syntax*. W3C Recommendation, World Wide Web Consortium. <http://www.w3.org/2007/OWL/draft/owl2-syntax/>
- Mouakher, A., Belkaroui, R., Bertaux, A., Labbani, O., Hugol-Gential, C., & Nicolle, C. (2019). An ontology-based monitoring system in Vineyards of the Burgundy Region. In *28th IEEE international conference on enabling technologies* (pp. 307–312). IEEE. <https://doi.org/10.1109/WETICE.2019.00070>
- Musen, M. A. (2015). The Protégé project: A look back and a look forward. *AI Matters*, 1(4), 4–12. <https://doi.org/10.1145/2757001.2757003>. <https://protege.stanford.edu/>
- Myklebust, E. B., Jiménez-Ruiz, E., Chen, J., Wolf, R., & Tollefsen, K. E. (2022). Prediction of Adverse Biological Effects of Chemicals Using Knowledge Graph Embeddings. *Semantic Web*, 13(3), 299–338. <https://doi.org/10.3233/SW-222804>

- Nakawala, H., Bianchi, R., Pescatori, L. E., Cobelli, O. D., Ferrigno, G., & Momi, E. D. (2019). “Deep-Onto” network for surgical workflow and context recognition. *International Journal of Computer Assisted Radiology Surgery*, 14(4), 685–696. <https://doi.org/10.1007/s11548-018-1882-8>
- Nardi, D., & Brachman, R. J. (2003). An introduction to description logics. In F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, & P. F. Patel-Schneider (Eds.), *The description logic handbook: Theory, implementation, and applications* (pp. 1–40). Cambridge University Press.
- Penov, Y., Piro, R., Motik, B., Horrocks, I., Wu, Z., & Banerjee, J. (2015). RDFox: A highly-scalable RDF store. In *The semantic web - 14th international conference, ISWC 2015, proceedings*, Lecture Notes in Computer Science (Vol. 9367, pp. 3–20). Springer Verlag. <https://www.oxfordsemantic.tech/product>; <https://ora.ox.ac.uk/objects/uuid:2a08b023-77be-431a-a08c-89b47381586a>
- Nickel, M., Murphy, K., Tresp, V., & Gabrilovich, E. (2015). A review of relational machine learning for knowledge graphs. *CoRR*. <http://arxiv.org/abs/1503.00759>
- Noy, N., & McGuinness, D. (2001). *Ontology development 101: A guide to creating your first ontology*, Technical Report, Knowledge Systems Laboratory, Stanford University. <http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf>
- Ontotext GraphDB. (2023). A popular RDF store. <https://www.ontotext.com/products/graphdb/>; [https://en.wikipedia.org/wiki/Ontotext\\_GraphDB](https://en.wikipedia.org/wiki/Ontotext_GraphDB)
- Peters, M. E., Neumann, M., IV, R. L. L., Schwartz, R., Joshi, V., Singh, S., & Smith, N. A. (2019). Knowledge enhanced contextual word representations. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 43–54). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-1005>
- Rossi, A., Barbosa, D., Firmani, D., Matinata, A., & Merialdo, P. (2021). Knowledge graph embedding for link prediction: A comparative analysis. *ACM Transactions on Knowledge Discovery from Data*, 15(2), 14:1–14:49. <https://doi.org/10.1145/3424672>
- Roy, K., Zi, Y., Narayanan, V., Gaur, M., & Sheth, A. P. (2023). Knowledge-infused self attention transformers. *CoRR*. <https://doi.org/10.48550/arXiv.2306.13501>
- Sarker, M. K., Zhou, L., Eberhart, A., & Hitzler, P. (2021). Neuro-symbolic artificial intelligence: Current trends. *CoRR*. <https://arxiv.org/abs/2105.05330>
- The Semantic Web Stack. (2022). [https://en.wikipedia.org/wiki/Semantic\\_Web\\_Stack](https://en.wikipedia.org/wiki/Semantic_Web_Stack)
- The Semantic Web Wiki. (2001). [https://www.w3.org/2001/sw/wiki/Main\\_Page](https://www.w3.org/2001/sw/wiki/Main_Page)
- Serafini, L., & d’Avila Garcez, A. S. (2016). Logic tensor networks: Deep learning and logical reasoning from data and knowledge. In *Proceedings of the 11th international workshop on neural-symbolic learning and reasoning (NeSy’16)* (Vol. 1768). CEUR Workshop Proceedings, CEUR-WS.org. [http://ceur-ws.org/Vol-1768/NESY16\\_paper3.pdf](http://ceur-ws.org/Vol-1768/NESY16_paper3.pdf)
- Sheth, A. P., Gaur, M., Kursuncu, U., & Wickramarachchi, R. (2019). Shades of knowledge-infused learning for enhancing deep learning. *IEEE Internet Computing*, 23(6), 54–63. <https://doi.org/10.1109/MIC.2019.2960071>
- Sheth, A. P., Roy, K., & Gaur, M. (2023). Neurosymbolic artificial intelligence (Why, What, and How). *IEEE Intelligent Systems*, 38(3), 56–62. <https://doi.org/10.1109/MIS.2023.3268724>
- Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., & Katz, Y. (2007). Pellet: A practical OWL-DL reasoner. *Journal of Web Semantics*, 5(2), 51–53. <https://doi.org/10.1016/j.websem.2007.03.004>
- Tanon, T. P., Weikum, G., & Suchanek, F. M. (2020). YAGO 4: A reasonable knowledge base. In *The semantic web - 17th international conference, ESWC 2020, proceedings* (Vol. 12123, pp. 583–596). Lecture Notes in Computer Science, Springer. [https://doi.org/10.1007/978-3-030-49461-2\\_34](https://doi.org/10.1007/978-3-030-49461-2_34)
- Trinh, T. H., Wu, Y., Le, Q. V., He, H., & Luong, T. (2024). Solving olympiad geometry without human demonstrations. *Nature*, 625, 476–482. <https://doi.org/10.1038/s41586-023-06747-5>
- Vrandečić, D., & Krötzsch, M. (2014). Wikidata: A free collaborative knowledgebase. *Communications of the ACM*, 57(10), 78–85.
- Wang, M., Qiu, L., & Wang, X. (2021). A survey on knowledge graph embeddings for link prediction. *Symmetry*, 13(3). <https://doi.org/10.3390/sym13030485>
- Whetzel, P. L., Noy, N. F., Shah, N. H., Alexander, P. R., Nyulas, C., Tudorache, T., & Musen, M. A. (2011). BioPortal: Enhanced functionality via new web services from the national center for biomedical ontology to access and use ontologies in software applications. *Nucleic Acids Research*, 39(suppl 2), W541–W545. <https://bioportal.bioontology.org/>
- Yamada, I., Asai, A., Shindo, H., Takeda, H., & Matsumoto, Y. (2020). LUKE: Deep contextualized entity representations with entity-aware self-attention. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 6442–6454). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.emnlp-main.523>
- Zhang, Z., Han, X., Liu, Z., Jiang, X., Sun, M., & Liu, Q. (2019). ERNIE: Enhanced language representation with informative entities. In *Proceedings of the conference of the association for computational linguistics* (pp. 1441–1451). Association for Computational Linguistics. <https://doi.org/10.18653/v1/p19-1139>