Empowering Knowledge Bases:
A Machine Learning Perspective
(Invited Paper and Talk)

Claudia d'Amato

Computer Science Department – University of Bari
`claudia.damato@uniba.it`

Abstract. The construction of Knowledge Bases requires quite often the intervention of knowledge engineering and domain experts, resulting in a time consuming task. Alternative approaches have been developed for building knowledge bases from existing sources of information such as web pages and crowdsourcing; seminal examples are NELL, DBPedia, YAGO and several others. With the goal of building very large sources of knowledge, as recently for the case of Knowledge Graphs, even more complex integration processes have been set up, involving multiple sources of information, human expert intervention, crowdsourcing. Despite significant efforts for making Knowledge Graphs as comprehensive and reliable as possible, they tend to suffer of incompleteness and noise, due to the complex building process. Nevertheless, even for highly human curated knowledge bases, cases of incompleteness can be found, for instance with disjointness axioms missing quite often. Machine learning methods have been proposed with the purpose of refining, enriching, completing and possibly raising potential issues in existing knowledge bases while showing the ability to cope with noise. The talk will concentrate on classes of mostly symbol-based machine learning methods, specifically focusing on concept learning, rule learning and disjointness axioms learning problems, showing how the developed methods can be exploited for enriching existing knowledge bases. During the talk it will be highlighted as, a key element of the illustrated solutions, is represented by the integration of: background knowledge, deductive reasoning and the evidence coming from the mass of the data. The last part of the talk will be devoted to the presentation of an approach for injecting background knowledge into numeric-based embedding models to be used for predictive tasks on Knowledge Graphs.[1]

1 Introduction

The construction of Knowledge Bases (KBs) quite often results a time consuming task since the intervention of knowledge engineering and domain experts is needed. For this reason, alternative solutions have been developed, exploiting

---

existing sources of information such as web pages; seminal examples are NELL[2], DBpedia[3], YAGO[4]. Recently the construction of very large KBs, named Knowledge Graphs (KGs), has been targeted and several examples already exist, spanning from enterprise products, such as those built by Google and Amazon to other open KGs such as Freebase[5] and Wikidata[6], to mention a few of them.

A KG is a graph of data intended to convey knowledge of the real world, conforming to a graph-based data model where nodes represent entities of interest and edges represent possibly different relations between these entities, furthermore the data graph can be enhanced with schema, generally by the use of ontologies that are employed to define and reason about the semantics of nodes and edges [24]. KGs are very large data collections requiring even more complex integration processes than in the past, involving multiple sources of information, human expert intervention, crowdsourcing. Despite significant efforts for making KGs as comprehensive as possible, it is well known [24] they tend to suffer of incompleteness and noise, due to the complex building process. Nevertheless, even for highly human curated KBs as for some ontologies, cases of incompleteness can be found, for instance with disjointness axioms missing quite often [43] that, as a consequence, may generate noise, that is invalid information with respect to the domain of reference despite of having a consistent KB.

Machine Learning (ML) methods, mostly grounded on inductive approaches, have been proposed for refining, enriching, completing and possibly raising issues in existing KBs while showing the ability to cope with noise [13]. Problems such as link prediction but also query answering and instance retrieval have been regarded as classification problems. Suitable methods, often inspired by symbol-based solutions in the *Inductive Logic Programming* (ILP) [14] field (aiming at inducing an hypothesised logic program from a background knowledge and a collection of examples), have been proposed [9, 27, 31, 19, 39]. Most of them are able to cope with expressive representation languages such as Description Logics (DLs) [2], theoretical foundation for OWL[7], standard representation language in the Semantic Web [4], and the *Open World Assumption* (OWA) typically adopted, differently from the *Closed Wold Assumption* (CWA) that is usually assumed in the traditional ML settings. Also, problems such as ontology refinement and enrichment at terminology/schema level, e.g. providing complex descriptions for a given concept name or assessing disjointness axioms, have been regarded as concept learning problems to be solved via supervised/unsupervised inductive learning methods for DL representations [17, 18, 30, 42, 36].

Nowadays, numeric-based ML methods such as *embeddings* [6, 26] and *deep learning* [15] solutions are receiving major attention because of their impressive ability to scale when applied to very large data collections. Mostly KG refine-

---

[2] http://rtw.ml.cmu.edu/rtw/
[3] https://www.dbpedia.org/
[4] https://yago-knowledge.org/
[5] https://developers.google.com/freebase
[6] https://www.wikidata.org/
[7] https://www.w3.org/OWL/

ment tasks, and specifically link/type predictions and triple classifications are targeted, with the goal of improving/limiting incompleteness in KGs. Nevertheless, the important gain, in terms of scalability, that numeric-based methods are obtaining is penalizing: a) the possibility to have interpretable models as a result of a learning process; b) the ability to exploit deductive (and complementary forms of) reasoning; c) the expressiveness of the representations to be considered and the compliance with the OWA.

For these reasons, the talk will focus primarily on the advances in the class of symbol-based ML methods, specifically analyzing concept learning, rule learning and disjointness axioms learning problems and related solutions to be used for enriching existing KBs. The key point is represented by the integration of: background knowledge, deductive reasoning and the evidence coming from the mass of data. The main idea consists in exploiting the evidence coming from assertional knowledge and deductive reasoning for drawing plausible conclusions to be possibly represented with intensional models. The last part of the talk will be dedicated to the presentation of an approach for injecting background knowledge into numeric-based ML methods and specifically embedding models to be used for predictive tasks on KGs.

Accordingly, in the following, concept learning (see Sect. 2), rule learning (see Sect. 3) and disjointness axioms learning (see Sect. 4) are briefly surveyed while semantically enriched numeric based solutions are illustrated in Sect. 5. Discussions on future research directions and conclusions are reported in Sect. 6.

2   Concept Learning for Ontology Enrichment

With the purpose of enriching ontologies at terminological level, methods for learning concept descriptions for a concept name have been proposed. The problem has been regarded as a supervised concept learning problem aiming at approximating an intensional DLs definition, given a set of individuals of an ontological KB acting as positive/negative training examples.

Various solutions, e.g. DL-Foil [17] and CELOE [30] (part of the DL-LEARNER suite[8]), have been formalized. They are mostly grounded on a *separate-and-conquer* (sequential covering) strategy: a new concept description is built by specializing, via suitable *refinement operators*, a partial solution to correctly cover (i.e. decide a consistent classification for) as many training instances as possible. Whilst DL-Foil works under OWA, CELOE works under CWA. Both of them may suffer of ending up in sub-optimal solutions. In order to overcome such issue, DL-Focl [37], Parcel [40] and SpACEL [41] have been proposed. DL-Focl is an optimized version of DL-Foil, implementing a base greedy covering strategy. Parcel combines top-down and bottom-up refinements in the search space. Specifically, the learning problem is split into various sub-problems, according to a divide-and-conquer strategy, that are solved by running CELOE as a subroutine. Once the partial solutions are obtained, they are combined in

---

[8] https://dl-learner.org/.

a bottom-up fashion. SPACEL extends PARCEL by performing a symmetrical specialization of a concept description.

These solutions proved to be able to learn approximated concept descriptions for a target concept name to be used for possibly introducing new (inclusion or equality) axioms in the KB, nevertheless, quite often, relatively small ontological KBs have been considered for the experiments, revealing that currently these solutions have limited ability to scale on very large KBs such as KGs.

3   Rule Learning for Knowledge Completion

Knowledge completion consists in finding new information at assertional level, that is facts, that are missing in a considered KB. This task has received increasing attention with the development of KGs that are well known to be incomplete, since it is also strongly related to the link prediction task (see Sect. 5).

One of the most well known system for knowledge completion of RDF[9] knowledge bases is AMIE [19]. Inspired by *association rule mining* [1] and ILP literature, AMIE aims at mining logic rules from RDF knowledge bases with the final goal of predicting new facts, that are RDF triples. AMIE (and its optimized version AMIE+ [20]) currently represents the most scalable rule mining system for learning rules on large RDF data collections and it is also explicitly tailored to support the OWA. However, it does not exploit any form of deductive reasoning.

A related rule mining system, similarly based on a level-wise generate and test strategy has been proposed [12]. It aims at learning SWRL rules [25] from OWL ontologies while exploiting schema level information and deductive reasoning during learning process. Both AMIE and the solution presented in [12] showed the ability to mine useful rules and to predict new assertional knowledge. However, the solution proposed in [12] showed limited ability to scale due to the exploitation of the reasoning capabilities. Nevertheless, with [12] additional rules could be obtained to be exploited for generalizing towards other kind of axioms (besides facts) such as general inclusion axioms, to be only validated by a domain expert or knowledge engineering.

4   Learning Disjointness Axioms

Disjointness axioms are essential for making explicit the negative knowledge about a domain, yet they are often overlooked when modelling ontologies [43] (thus also affecting the efficacy of reasoning services). To tackle this problem, automated methods for discovering these axioms from the data distribution have been devised.

A solution grounded on *association rule mining* [1] has been proposed in [42]. It is based on studying the correlation between classes comparatively, namely *association rules*, *negative association rules* and *correlation coefficient*. Background knowledge and reasoning capabilities are used to a limited extent. Different approaches have been instead proposed in [31] and [3] where *relational*

---

[9] https://www.w3.org/RDF/

learning methods and techniques based on *formal concept analysis* have been respectively employed to the purpose. However, no specific assessment of the quality of the induced axioms is made.

A different solution has been proposed in [36] where, moving from the assumption that two or more concepts may be mutually disjoint when the sets of their (known) instances do not *overlap*, the problem has been regarded as a clustering problem, aiming at finding partitions of similar individuals of the KB, according to a *cohesion* criterion quantifying the degree of homogeneity of the individuals in an element of the partition. Specifically, the problem has been cast as a *conceptual clustering* problem, where the goal is both to find the best possible partitioning of the individuals and also to induce intensional definitions of the corresponding classes expressed in the standard representation languages. Emerging disjointness axioms are captured by the employment of *terminological cluster trees* (TCTs) and by minimizing the risk of mutual overlap between concepts. Once the TCT is grown, groups of (disjoint) clusters located at sibling nodes identify concepts involved in candidate disjointness axioms to be derived. Unlike [42], based on the statistical correlation between instances, the empirical evaluation of [36] showed its ability to discover disjointness axioms also involving complex concept descriptions, thanks to the exploitation of the underlying ontology as background knowledge.

5   Enriched Embedding Models for Knowledge Graph Completion

Symbol-based ML methods adopt symbols for representing entities and relationships of a domain and infer generalizations, ideally readily interpretable even by the humans, that provide new insights into the data. Differently, numeric-based methods typically adopt feature vector (propositional) representations and cannot provide interpretable models but they usually result rather scalable. For this reason numeric-based solutions have been mainly used for performing *link prediction* in KGs (also referred to as *knowledge graph completion*), which amounts to predict the existence (or the probability of correctness) of triples in the large KGs, that are known to be often missing facts. Mostly RDF representation language has been targeted and almost no reasoning is exploited; most expressive languages (such as OWL) are basically discarded.

Among the others, KG embedding methods have received the greatest attention. They aim at converting the data graph into an optimal low-dimensional space in which *graph structural information* and *graph properties* are preserved as much as possible [6, 26]. The low-dimensional spaces enable computationally efficient solutions that scale better with the KG dimensions. Graph embedding methods may differ in their main building blocks: the *representation space* (e.g. point-wise, complex, discrete, Gaussian, manifold), the *encoding model* (e.g. linear, factorization, neural models) and the *scoring function* (that can be based on distance, energy, semantic matching or other criteria) [26]. In any case, the

objective consists in learning embeddings such that the score of a valid (positive) triple is lower than the score of an invalid (negative) triple.

TRANSE [5] has been the very first embedding model registering very high scalability performances. The method relies on a *stochastic optimization process*, that iteratively updates the distributed representations by increasing the score of the positive triples i.e. the observed triples, while lowering the score of un-observed triples standing as negative examples. The embedding of all entities and predicates in the KG is learned by minimizing a *margin-based ranking loss*. Nevertheless, TRANSE resulted limited in representing properly various types of properties such as *reflexive* ones, and 1-to-N, N-to-1 and N-to-N relations. To tackle this limitation, moving from TRANSE, a large family of models has been originated; among the others, TRANSR [32], has been proposed resulting as more suitable to handle non 1-to-1 relations.

An important point that needs to be highlighted is that, because RDF representation is mostly tackled, most of the considered data collections only contain positive (training) examples, since usually false facts are not encoded. However, as negative examples are needed when learning vector embeddings, for obtaining negative examples two different approaches are generally adopted: either *corrupting* true/observed triples with the goal of generating plausible negative examples or making a *local-closed world assumption* (LCWA) in which the data collection is assumed as *locally* complete [35]. In both cases, wrong negative information may be generated and thus used when training and learning the embedding models. Even more so, existing embedding models do not make use of the additional semantic information that may be coded when more expressive representation languages are adopted.

Recently, empowered embedding models, targeting KG refinement tasks, have been proposed. In [22] a KG embedding method considering also logical rules has been formalized, where triples in the KG and rules are represented in a unified framework. A common loss over both representations is defined which is minimized to learn the embeddings. This proposal resulted in a novel solution but the specific form of prior knowledge that needs to be available for the KG constitutes its main drawback. A similar drawback also applies to [34], where a solution based on adversarial training is formalized, exploiting Datalog clauses to encode assumptions which are used to regularize neural link predictors. An inconsistency loss is derived that measures the degree of violation of such assumptions on a set of adversarial examples.

Complementary solutions have been developed. Besides graph structural information, they exploit the additional knowledge that is typically available when rich representation languages as RDFS and OWL are employed. Differently from the works previously mentioned, these solutions do not require any specific additional formalisms for representing prior knowledge. Particularly, in [33] it has been proved the effectiveness of combining embedding methods and strategies relying on reasoning services for the injection of background knowledge to enhance the performance of a specific predictive model. Following this line, TRANsOWL, aiming at injecting background knowledge during the learning process,

and its upgraded version TRANSROWL, where a newly defined and more suitable loss function and scoring function are also exploited, have been proposed [11, 10]. These solutions can take advantage of an informed corruption process that leverages on reasoning capabilities, while limiting the amount of false negatives that a less informed random corruption process may cause. Specifically, TRANSOWL formalizes a model characterized by two main components devoted to inject background knowledge in the embedding-based model during the training phase: 1) *Reasoning:* It is used for generating corrupted triples that can certainly represent negative instances, thus avoiding false negatives, for a more effective model training. Precisely, using a reasoner, corrupted triples are generated exploiting available RDF/OWL axioms, particularly domain, range, disjointWith, functionalProperty; 2) *Background Knowledge Injection:* A set of different axioms, that are equivalentClass, equivalentProperty, inverseOf and subClassOf, are employed for defining constraints on the score function to be used in the training phase, so that the resulting vectors, related to such axioms, reflect their specific properties. As a consequence, new triples are also added to the training set on the grounds of these axioms.

These models have been proved improving their effectiveness on link prediction and triple classification tasks when compared to models such as TRANSE and TRANSR, grounded on the use of a random corruption process and structural graph properties only. Nevertheless, some shortcomings also emerged, precisely when some of the considered schema axioms were missing, suggesting that additional efforts need to be pursued in this direction.

6   Discussion and Conclusions

Existing symbol-based ML methods are not actually comparable, in terms of scalability, to numeric-based ML solutions. Nevertheless, as already discussed in Sect. 1 and Sect. 5, this gain is not for free since it is obtained mostly by giving up expressive representation languages, such as OWL, and by almost forgetting one of the most powerful characteristic of these languages, that is being empowered with deductive reasoning capabilities that allow to derive new knowledge. Furthermore, differently from symbol-based methods, numeric-based solutions lack of the ability to provide interpretable models, thus limiting the possibility to interpret and understand the motivations for the returned results. Additionally, tasks such as concept and/or disjointness axioms learning cannot be performed without symbol-based methods which can certainly benefit of very large amount of information for providing potentially more accurate results.

Even if some initial results have been registered for semantically enriched embedding solutions (Sect. 5), significant research efforts need to be devoted towards developing ML solutions that, while keeping scalability, are able to target expressive representations as well as to provide interpretable models. This actually means pushing towards the integration of numeric and symbolic approaches. Some discussions in this direction have been developed by the Neural-Symbolic Learning and Reasoning community [21, 23], which seeks to integrate principles

from neural networks learning and logical reasoning. The main conclusion has been that neural-symbolic integration appears particularly suitable for applications characterized by the joint availability of large amounts of (heterogeneous) data and knowledge descriptions, which is actually the case for KGs. Additionally, a set of key challenges and opportunities have been outlined [21], such as: how to represent expressive logics within neural networks, how neural networks should reason with variables, or how to extract symbolic representation from trained neural networks. Preliminary results have been recently registered, encouraging pursuing the research direction. An example is represented by SimplE [28], a scalable tensor-based factorization model that is able to learn interpretable embeddings incorporating logical rules through weight tying. Ideas for extracting propositional rules from trained neural networks under a background knowledge have been illustrated in [29], showing that the exploitation of a background knowledge allows: a) to reduce the cardinality of extracted rule sets; b) to reproduce the input-output function of the trained neural network.

A conceptual sketch for explaining the classification behavior of artificial neural networks in a non-propositional setting with the use of a background knowledge has been proposed [38]. This sheds the light on another important issue, that is, the necessity of providing explanations to ML results [8], particularly when they come from very large KBs. The solution illustrated in [38] is in agreement with the idea of exploiting symbol-based interpretable models to explain conclusions [14]. Nevertheless, interpretable models describe *how* solutions are obtained but not *why* they are obtained, whilst, as argued in [16, 21], providing an explanation should mean supplying a line of reasoning illustrating the decision making process of a model using human understandable features [7]. Hence, in a more broad sense, providing an explanation requires to open the box of the reasoning process and make it understandable.

References

1. Agrawal, R., Imieliński, T., Swami, A.: Mining association rules between sets of items in large databases. In: Buneman, P., et al. (eds.) Proc. of the 1993 ACM SIGMOD Int. Conf. on Management of Data, pp. 207–216. ACM (1993). https://doi.org/10.1145/170035.170072
2. Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (eds.): Description Logic Handbook, 2nd edition. Cambridge University Press (2010). https://doi.org/10.1017/CBO9780511711787
3. Baader, F., Ganter, B., Sertkaya, B., Sattler, U.: Completing description logic knowledge bases using formal concept analysis. In: Veloso, M. (ed.) IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence. pp. 230–235. AAAI Press (2007)
4. Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Scientific American **284**(5), 34–43 (2001)
5. Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., Yakhnenko, O.: Translating embeddings for modeling multi-relational data. In: Burges, C.J.C., et al. (eds.) Proceedings of NIPS 2013, pp. 2787–2795. Curran Associates, Inc. (2013)

6. Cai, H., Zheng, V.W., Chang, K.: A comprehensive survey of graph embedding: Problems, techniques, and applications. IEEE Transactions on Knowledge and Data Engineering **30**(09), 1616–1637 (2018). https://doi.org/10.1109/TKDE.2018.2807452

7. Chen, J., Lécué, F., Pan, J., Horrocks, I., Chen, H.: Knowledge-based transfer learning explanation. In: Thielscher, M., et al. (eds.) Principles of Knowledge Representation and Reasoning: Proc. of the 16th International Conference, KR 2018. pp. 349–358. AAAI Press (2018)

8. d'Amato, C.: Logic and learning: Can we provide explanations in the current knowledge lake? In: Bonatti, P., et al. (eds.) Knowledge Graphs: New Directions for Knowledge Representation on the Semantic Web (Dagstuhl Seminar 18371), Dagstuhl Reports, vol. 8, pp. 37–38. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik (2019). https://doi.org/10.4230/DagRep.8.9.29

9. d'Amato, C., Fanizzi, N., Esposito, F.: Query answering and ontology population: An inductive approach. In: Bechhofer, S., et al. (eds.) The Semantic Web: Research and Applications, 5th European Semantic Web Conference, ESWC 2008, Proceedings. LNCS, vol. 5021, pp. 288–302. Springer (2008). https://doi.org/10.1007/978-3-540-68234-9_23

10. d'Amato, C., Quatraro, N.F., Fanizzi, N.: Embedding models for knowledge graphs induced by clusters of relations and background knowledge. In: 1st International Joint Conf. on Learning & Reasoning, IJCLR 2021, Proceedings (2021), to appear

11. d'Amato, C., Quatraro, N.F., Fanizzi, N.: Injecting background knowledge into embedding models for predictive tasks on knowledge graphs. In: Verborgh, R., et al. (eds.) The Semantic Web International Conference, ESWC 2021, Proceedings. LNCS, vol. 12731, pp. 441–457. Springer (2021). https://doi.org/10.1007/978-3-030-77385-4_26

12. d'Amato, C., Tettamanzi, A., Tran, D.M.: Evolutionary discovery of multi-relational association rules from ontological knowledge bases. In: Blomqvist, E., et al. (eds.) Knowledge Engineering and Knowledge Management International Conference, EKAW 2016, Proceedings. LNCS, vol. 10024, pp. 113–128 (2016). https://doi.org/10.1007/978-3-319-49004-5_8

13. d'Amato, C.: Machine learning for the semantic web: Lessons learnt and next research directions. Semantic Web **11**(1), 195–203 (2020). https://doi.org/10.3233/SW-200388

14. De Raedt, L. (ed.): Logical and Relational Learning: From ILP to MRDM (Cognitive Technologies). Springer-Verlag (2008)

15. Deng, L., Yu, D. (eds.): Deep Learning: Methods and Applications. NOW Publishers (2014). https://doi.org/10.1561/2000000039

16. Doran, D., Schulz, S., Besold, T.: What does explainable AI really mean? A new conceptualization of perspectives. In: Besold, T.R., Kutz, O. (eds.) Proceedings of the 1st International Workshop on Comprehensibility and Explanation in AI and ML 2017 co-located with 16th Int. Conf. of the Italian Association for Artificial Intelligence (AI*IA 2017), CEUR Work. Proc., vol. 2071. CEUR-WS.org (2017)

17. Fanizzi, N., d'Amato, C., Esposito, F.: DL-FOIL concept learning in description logics. In: Zelezný, F., Lavrac, N. (eds.) Inductive Logic Programming, 18th International Conference, ILP 2008, Proceedings. LNCS, vol. 5194, pp. 107–121. Springer (2008). https://doi.org/10.1007/978-3-540-85928-4_12

18. Fanizzi, N., d'Amato, C., Esposito, F.: Metric-based stochastic conceptual clustering for ontologies. Inf. Syst. **34**(8), 792–806 (2009). https://doi.org/10.1016/j.is.2009.03.008

19. Galárraga, L., Teflioudi, C., Hose, K., Suchanek, F.M.: AMIE: association rule mining under incomplete evidence in ontological knowledge bases. In: Schwabe, D., et al. (eds.) 22nd International World Wide Web Conference, WWW '13. pp. 413–422. International World Wide Web Conferences Steering Committee / ACM (2013). https://doi.org/10.1145/2488388.2488425
20. Galárraga, L., Teflioudi, C., Hose, K., Suchanek, F.M.: Fast rule mining in ontological knowledge bases with AMIE+. VLDB Journal **24**(6), 707–730 (2015). https://doi.org/10.1007/s00778-015-0394-1
21. d'Avila Garcez, A., Besold, T., De Raedt, L., Földiák, P., Hitzler, P., Icard, T., Kühnberger, K., Lamb, L., Miikkulainen, R., Silver, D.: Neural-symbolic learning and reasoning: Contributions and challenges. In: 2015 AAAI Spring Symposia. AAAI Press (2015), `http://www.aaai.org/ocs/index.php/SSS/SSS15/paper/view/10281`
22. Guo, S., Wang, Q., Wang, L., Wang, B., Guo, L.: Jointly embedding knowledge graphs and logical rules. In: Proceedings of EMNLP 2016. pp. 192–202. ACL (2016). https://doi.org/10.18653/v1/D16-1019
23. Hitzler, P., Bianchi, F., Ebrahimi, M., Sarker, M.K.: Neural-symbolic integration and the semantic web. Semantic Web Journal **11**(1), 3–11 (2020). https://doi.org/10.3233/SW-190368
24. Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., de Melo, G., Gutiérrez, C., Gayo, J.L., Kirrane, S., Neumaier, S., Polleres, A., Navigli, R., Ngomo, A.N., Rashid, S., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., Zimmermann, A.: Knowledge graphs. ACM Computing Surveys **54**, 1–37 (2021). https://doi.org/https://doi.org/10.1145/3447772
25. Horrocks, I., Patel-Schneider, P., Boley, H., Tabet, S., Grosof, B., Dean., M.: SWRL: A semantic web rule language combining owl and ruleml (2004), `http://www.aaai.org/ocs/index.php/SSS/SSS15/paper/view/10281`
26. Ji, S., Pan, S., Cambria, E., Marttinen, P., Yu, P.S.: A survey on knowledge graphs: Representation, acquisition and applications arXiv:2002.00388 (2020)
27. Józefowska, J., Lawrynowicz, A., Lukaszewski, T.: The role of semantics in mining frequent patterns from knowledge bases in description logics with rules. TPLP **10**(3), 251–289 (2010). https://doi.org/10.1017/S1471068410000098
28. Kazemi, S., Poole, D.: Simple embedding for link prediction in knowledge graphs. In: Bengio, S., et al. (eds.) Advances in Neural Information Processing Systems 31: Annual Conf. on Neural Information Processing Systems 2018, NeurIPS 2018, pp. 4289–4300. ACM (2018)
29. Labaf, M., Hitzler, P., Evans, A.: Propositional rule extraction from neural networks under background knowledge. In: Besold, T.R., et al. (eds.) Proceedings of the Twelfth International Workshop on Neural-Symbolic Learning and Reasoning, NeSy 2017. CEUR Workshop Proceedings, vol. 2003. CEUR-WS.org (2017)
30. Lehmann, J., Auer, S., Bühmann, L., Tramp, S.: Class expression learning for ontology engineering. J. Web Semant. **9**(1), 71–81 (2011). https://doi.org/10.1016/j.websem.2011.01.001
31. Lehmann, J., Bühmann, L.: ORE - a tool for repairing and enriching knowledge bases. In: Patel-Schneider, P.F., et al. (eds.) The Semantic Web - ISWC 2010 - 9th International Semantic Web Conference, ISWC 2010, Revised Selected Papers, Part II. LNCS, vol. 6497, pp. 177–193. Springer (2010). https://doi.org/10.1007/978-3-642-17749-1_12
32. Lin, Y., Liu, Z., Sun, M., Liu, Y., Zhu, X.: Learning entity and relation embeddings for knowledge graph completion. In: AAAI 2015 Proceedings. p. 2181–2187. AAAI Press (2015)

33. Minervini, P., Costabello, L., Muñoz, E., Novácek, V., Vandenbussche, P.: Regularizing knowledge graph embeddings via equivalence and inversion axioms. In: Ceci, M., et al. (eds.) Proceedings of ECML PKDD 2017, Part I. LNAI, vol. 10534, pp. 668–683. Springer (2017). https://doi.org/10.1007/978-3-319-71249-9_40

34. Minervini, P., Demeester, T., Rocktäschel, T., Riedel, S.: Adversarial sets for regularising neural link predictors. In: Elidan, G., et al. (eds.) UAI 2017 Proceedings. AUAI Press (2017)

35. Nickel, M., Murphy, K., Tresp, V., Gabrilovich, E.: A review of relational machine learning for knowledge graphs. Proceedings of the IEEE **104**(1), 11–33 (2016). https://doi.org/10.1109/JPROC.2015.2483592

36. Rizzo, G., d'Amato, C., Fanizzi, N., Esposito, F.: Terminological cluster trees for disjointness axiom discovery. In: Blomqvist, E., et al. (eds.) The Semantic Web - 14th International Conference, ESWC 2017, Proceedings, Part I. LNCS, vol. 10249, pp. 184–201. Springer (2017). https://doi.org/10.1007/978-3-319-58068-5_12

37. Rizzo, G., Fanizzi, N., d'Amato, C., Esposito, F.: A framework for tackling myopia in concept learning on the web of data. In: Faron-Zucker, C., et al. (eds.) Knowledge Engineering and Knowledge Management International Conference, EKAW 2018, Proceedings. LNCS, vol. 11313, pp. 338–354. Springer (2018). https://doi.org/10.1007/978-3-030-03667-6_22

38. Sarker, M., Xie, N., Doran, D., Raymer, M., Hitzler, P.: Explaining trained neural networks with semantic web technologies: First steps. In: Besold, T.R., et al. (eds.) Proceedings of the 12th Int. Workshop on Neural-Symbolic Learning and Reasoning, NeSy 2017. CEUR Workshop Proc., vol. 2003. CEUR-WS.org (2017)

39. Tiddi, I., d'Aquin, M., Motta, E.: Dedalo: Looking for clusters explanations in a labyrinth of linked data. In: Presutti, V., et al. (eds.) The Semantic Web: Trends and Challenges - 11th Int. Conference, ESWC 2014, Proceedings. LNCS, vol. 8465, pp. 333–348. Springer (2014). https://doi.org/10.1007/978-3-319-07443-6_23

40. Tran, A., Dietrich, J., Guesgen, H., Marsland, S.: An approach to parallel class expression learning. In: Bikakis, A., Giurca, A. (eds.) Rules on the Web: Research and Applications - 6th Int. Symposium, RuleML 2012, Proc., LNCS, vol. 7438, pp. 302–316. Springer (2012). https://doi.org/10.1007/978-3-642-32689-9_25

41. Tran, A., Dietrich, J., Guesgen, H., Marsland, S.: Parallel symmetric class expression learning. J. of Machine Learning Research **18**(64), 1–34 (2017)

42. Völker, J., Fleischhacker, D., Stuckenschmidt, H.: Automatic acquisition of class disjointness. Journal of Web Semantics **35**(P2), 124–139 (2015). https://doi.org/10.1016/j.websem.2015.07.001

43. Wang, T.D., Parsia, B., Hendler, J.: A survey of the web ontology landscape. In: Cruz, I., et al. (eds.) The Semantic Web - ISWC 2006, 5th Int. Semantic Web Conference Proceedings. LNCS, vol. 4273. Springer (2006), doi: 10.1007/11926078_49