Exploiting Semantics for Explaining Link Prediction over Knowledge Graphs

Yashrajsinh Chudasama* [0000-0003-3422-366X]

Leibniz University, Hannover, Germany
yashrajsinh.chudasama@stud.uni-hannover.de

Abstract. The use of Symbolic and sub-symbolic AI techniques on Knowledge Graphs (*KGs*) has shown significant progress in several applications. However, many of these methods remain opaque, and the decision-making process behind them can be perplexing. This can result in a lack of trust and reliability in the overall framework. While various explainable frameworks have been proposed to address these issues, do not always provide a complete understanding and may raise privacy concerns as sensitive data may be revealed during the explanation process. In contrast, our proposed approach leverages the semantics of *KGs* and causal relationships to enhance explainability while still maintaining a high level of trust and reliability. By focusing on XAI for link prediction models and considering entailment regimes (e.g., `rdfs:subPropertyOf`), the approach can provide more comprehensive and accurate explanations. Moreover, the use of symbolic reasoning allows for more transparent and interpretable explanations. The preliminary results show that our approach is capable of exploiting the semantics of an entity in *KG* and enhancing the explanations. Henceforth, more work needs to be conducted, to fully comprehend all impacting factors and to identify the most relevant explanations of the machine learning models over *KGs*.

Keywords: Knowledge Graphs · Link Prediction · Explainability

1 Introduction

Recent advances in Artificial Intelligence (AI) have already started to impact our daily lives in terms of intelligence, and demonstrated their success in forecasting machine learning problems (e.g., disease diagnosis [9]). Explainability refers to the degree to which humans can understand the decisions made by computational frameworks. Extracting explanations is crucial, particularly because they are often obscure, and the explainability of the outcomes is partially achieved. Explainable predictive models have rapidly become a pertinent problem [4] in data management. Various approaches [6,11] attempt to understand the algorithmic decisions made by machine learning models, but they are unable to capture the insights of the model behavior to translate them into the domain.

* TIB Leibniz Information Centre for Science and Technology, Hannover, Germany
Category: Early Stage Ph.D.

KGs are data structures that encode data and knowledge together with domain ontologies representing real-world information, where entities like `Louis XIV` and `Marie Theresa` are linked via directed edges denoted by binary relationships forming triples called facts, i.e., ⟨Louis XIV, spouse, Marie Theresa⟩. In recent years, *KGs* have been built in various domains, and have led to a broad range of applications, including Knowledge Graph Completion (*KGC*) [1], or Query Processing [12]. *KGC* can discover new knowledge based on existing ones and check knowledge consistency. It is an appealing research topic that is important for completing and cleaning up *KGs*.
KGs represent knowledge in the form of factual statements of the form ⟨head entity, relation, tail entity⟩, shortened as ⟨e_h, r, e_t⟩. In the literature about *Knowledge Graph Embeddings (KGE)*, these notations are used to represent the facts in *KG*. *KGE* models, e.g., TransE, learn latent representations of entities and relations in continuous vector spaces, called embeddings, to preserve the *KG* structure. The most common learning methods for *KGC* are link predictions or triple classifications tasks based on a *KGE* model. Link Prediction (LP) confronts the issue of incompleteness by analyzing the already known facts to deduce new missing facts. For example, knowing the facts ⟨`Louis XIV, hasChild, Wessex`⟩ and ⟨`Wessex, hasMother, Marie Theresa`⟩, a LP model could predict ⟨`Louis XIV, spouse, Marie Theresa`⟩. However, these latent vector representations of the entities and the relations are not self-explainable, and an evaluation of the inductive abilities is still an open research issue.
Recently, the problem of explainable methods for link prediction has received attention [13,15]. Following the taxonomy by Rossi et al. [13] the *necessary* and *sufficient* explanations can be seen as either the set of facts in absence of which the link prediction model would not be able to yield the prediction; or a set of facts if given to an entity would lead the model to yield that prediction. For instance, given a tail prediction ⟨`Berlin, country, Germany`⟩, the facts about head entity `Berlin`: ⟨`Berlin, capital, Germany`⟩, and ⟨`Berlin, located, Germany`⟩ if removed from the training facts, leads the model to change the predicted tail. Thus, these facts were the necessary for the model to predict the correct tail entity, i.e., `Germany` with relation `country`. In *sufficient* explanation scenario, for example, when explaining the tail prediction ⟨`Berlin, country, Germany`⟩, identifying all the training facts about `Berlin`, if given to any head entity in the training facts, can lead the model to predict their country as Germany. For instance, adding the fact ⟨`Washington D.C., capital, Germany`⟩ to the training model is enough to yield the predicted country for `Washington D.C.` to be `Germany`. One of the crucial tasks for embedding-based explanation is efficiently learning and extracting explanations not only considering the data graphs, but also the meaning of the data given an ontology. When there are more triples or relations to consider, embedding-based reasoning is more effective.
SHACL[1] (the Shapes Constraint Language) is the W3C recommendation for defining integrity constraints over knowledge graphs. To trace and enhance the explanation for the predictive models built over data collected from *KGs*, our

[1] https://www.w3.org/TR/2017/REC-shacl-20170720/

approach **InterpretME**[2,3] relies on a symbolic system, currently, this system validates integrity constraints that provide a meaningful description of an entity of a prediction model. The current version of **InterpretME** is customized for supervised machine learning models (e.g., Decision Trees), embedding models (e.g., TransE), and interpretable tools (e.g., LIME [11]). In this proposal, the approach of **InterpretME** is introduced to fill the gap towards the application of LP Explainability over the *KGs*. The preliminary results of the research reveal the key role of Semantic Web technologies in explainable AI and demonstrate the importance of considering entailment regimes for the extraction of explanations.

1.1 Motivating Example

The motivation of our work originates from the lack of explainability methods with machine learning models over *KGs*. Although state-of-art techniques provide automated machine-learning pipelines, they are unable to generate human- and machine-readable decisions to assist users and enhance their efficiency. In this proposal, explainability over the link prediction tasks is considered as an application. This task can be subdivided into a tail prediction task, which predicts the most plausible tail e_t and a head prediction task that predicts the most plausible head e_h. Fig 1 depicts a naive approach that explains a link prediction task,

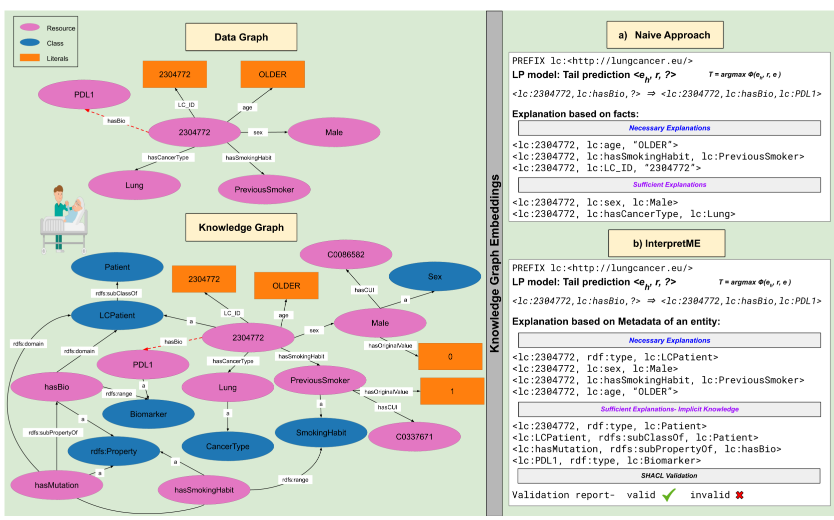

Fig. 1: **Motivating Example.** Explainable link prediction. a) Naive Approaches show facts with the head entity which leads to the particular prediction of a tail entity with *hasBio*. b) InterpretME depicts facts with the head entity and implicit knowledge which lead to the tail entity with *hasBio* and also provides SHACL validation reports.

[2] https://github.com/SDM-TIB/InterpretME
[3] https://github.com/SDM-TIB/InterpretME_Demo_ESWC2023

i.e., ⟨`lc:2304772, lc:hasBio, ?`⟩; the expected tail to be inferred is lc:PDL1. Embedding models (e.g., TransE) for link prediction (e.g., tail prediction) are utilized; they are executed on top of facts in data graphs.
Fig 1 illustrates the explanations based on link prediction tasks considering the RDF graph: **i)** as factual statements and, **ii)** with the meaning of an entity (i.e., `rdf:type, rdfs:domain, rdfs:subPropertyOf`). An input is collected from an **RDF *KG*** accessible via SPARQL endpoint, that integrates data about lung cancer patients. An RDF graph includes features describing the main characteristics of a lung cancer patient, i.e., patient identifier (a.k.a. LC_ID), gender, age, smoking habits, and lung cancer biomarkers. The predictive task is a link prediction of a tail entity into a low dimensional latent vector space to predict new infer facts about the patient by considering the neighborhood.
InterpretME resorts to Pykeen optimizer recommendations for hyperparameter optimization in the *KGE* models. Further, in the naive approach, the explainable tool Kelpie [13] is utilized to provide local interpretations of each patient in the training triples. Kelpie yields the relevant facts, by worsening or improving the scores of prediction, categorizing them into two categories *necessary* and *sufficient*. The terms *necessary* and *sufficient* are complementary to each other. Fig. 1 depicts an exemplar entity where Kelpie determines the most plausible explanations based on facts for the tail prediction task ⟨`lc:2304772, lc:hasBio, lc:PDL1`⟩ and generates the necessary explanations of a particular tail prediction are: ⟨`lc:2304772, lc:hasSmokingHabit, lc:PreviousSmoker`⟩, ⟨`lc:2304772, lc:age, "OLDER"`⟩. The naive approach outcomes allow for understanding the quality of the implemented framework. Although our user would have been able to understand the explanations generated by Kelpie, this user would need to trace these results back to the original data attributes to discover, for instance, whether the reported patient `lc:2304772` violates the domain integrity constraints or not. In contrast, **InterpetME** yields the explanations of the link prediction model based on the facts, implicit knowledge (e.g., `rdfs:subClassOf, rdfs:subPropertyOf`) and SHACL constraints to ensure the trustability of RDF data and predictions are consistent with the domain constraints.

2 State of the Art

The necessity of automated machine learning frameworks with assistance has gained tremendous popularity in various domains. Amongst the explainability over the *KGs*, the works most related to ours form four main categorizations: *XAI, Link Prediction, SHACL Validation, and Causal Models.*
XAI Frameworks. Within the Explainable AI community, there has been a surge of research on explainability techniques. These techniques are of two main categories: 1) Intrinsic, and 2) Post-hoc explainability. Intrinsic explainability refers to the machine learning models that are considered explainable due to their simple structure, e.g., decision trees. Post-hoc explainability embodies the fully trained black-box models, thus, trying to explain and justify the logic behind the model outputs. The advantage of such techniques is that they are model-

agnostic (i.e., the explanations can be generated across any models); LIME [11] is one of the exemplary post-hoc explainable methods. LIME aims to approximate any data-driven algorithm, with a local interpretable model to explain each instance prediction. Such techniques started showing major growth in many domains (e.g., Biomedical). However, this *saliency* explanations are human intuition matching techniques for entities and cannot be translated into a domain application. Our approach overcomes these limitations and provides fine-grained explanations linked to the entities in the*KGs*.

Link Prediction. *KGE* encodes the structure of triples in a KG, and can thus be used to perform link predictions, i.e., inferring the missing facts. Borrego et al. [3] propose a *KGC* approach using a set of neighborhood-aware features. However, the problem of learning embeddings for *KGs* has gained considerable attention, and only a few works address the explainability issues in link prediction over *KGs*. Zhang et al. [15] introduce a method of data poisoning; given a prediction $\langle e_h, r, e_t \rangle$, this data poisoning method identifies facts that, if are removed or added to training samples, they make worse the scoring function $\phi(e_h,\ r,\ e_t)$. Rossi et al. [13] propose the Kelpie framework, which explains the predicted links based on embedding via *necessary* and *sufficient* explanations. Rossi et al. state that the Kelpie framework computes the subset of training triples which can be seen as either the set of triples in addition or removal of which the model would yield that prediction. This framework is based on the aforementioned Post-hoc explainability. Nonetheless, these methods still lack in considering the semantic meaning of an entity and properties in a *KG*. Our approach aims at explaining the predictions based on the entailment regimes.

SHACL Validation. Explainability refers to the ability to interpret, understand and provide justifications for the decisions made by the machine learning models. In the context of *KGs*, SHACL validations are used to justify the machine learning model's prediction. Hence, defining the constraints on the structure and the data in the knowledge graph, SHACL is used to ensure that the predictions made by the link prediction model are consistent with the constraints and can provide justification for the predictions. The proposed approach relies on Trav-SHACL [5], Figuera et al. describe the capability of validating the shape schema against a SPARQL endpoint and scales better compared to other baseline approaches. Rohde et al. [12] report the perception of incorporating the SHACL validation result into SPARQL query answers by running the validation during the query processing. These validation results can provide one more layer of explainability. Thus, SHACL is more evident for enabling explainable AI in the context of *KGs*, provides validation, and explain the predictions.

Causal Models. A growing literature on causal models for the explainability of black-box models emphasizes that explanation is a normative goal that relates to real-world relationships (cause-effect) [2]. Pearl et al. [10] describe the essential role of the causal models via seven pillars which are beyond the reach of current machine learning models. In some recent work on the relational database, Salimi et al. [14] propose a declarative language, *CARL* which represents complex causal models using Horn clauses and constructs a unit table specific to the query and implements a causal model to identify the impact of treatment variable. In some of the closely related work, Huang [8] proposes *CareKG*, a causal query framework over the *KGs* to analyze the impact of treatment variables on the outcome and defines the aggregation function for multiple treatment variables. However, these approaches describe the formalism for the causal model but do not explain *"Why this particular decision?"* and also ignores the meaning of an entity in *KG*. To our best knowledge, none applied causality to provide more expressive explanations over *KGs*. Henceforth, one research focus of the proposal is to connect explainability and cause-effect analysis over *KGs*, so that the framework can provide more accurate explanations for its predictions and will greatly impact the Semantic Web community.

3 Problem Statement

Consider an RDF knowledge graph *KG* $(\zeta,\ R,\ G)$, defined as a directed edge labeled graph such that each node $e \in \zeta$ represents an entity, each $r \in R$ represents a unique relation, and each directed edge $\langle e_h, r, e_t \rangle \in G$ represents a fact about the head entity e_h. Given a tail prediction[4] $\langle e_h,\ r,\ ?\ \rangle$ where e_h is the head entity, r is the relation between entities, and $\langle e_h,\ r,\ e_t \rangle \notin G$, the aim is to find a set of most plausible entities e_t by inferring new facts based on the existing relations and entities in G and provides an interpretable set of facts $T = \{(e_h, \text{r}, \text{e}) | \text{e} \in \zeta\}$ which lead the black-box model to predict new facts. Unlike the previous method uses a fact-based approach to provide explanations. Our approach considers entailment regimes (i.e., `RDFS` and `OWL`) over *KGs* and causal relations; scalability is also one of our goals.

In contrast to baselines, the goal is to develop a framework that can be qualitatively and quantitatively evaluated for the explainability of machine learning models over *KGs*. To this end, we evaluate our approach by answering the following research questions: **RQ1)** What is the impact of integrating machine learning frameworks with KGs to enhance explainability? **RQ2)** To what extent do the extracted explanations comply with the predictions? **RQ3)** What is the impact of injecting RDFS and OWL axioms in the explanations over the *KGs*? With the above research questions, we aim to contribute to the Semantic Web and AI communities and develop a generalized framework that leverages black-box models to provide meaningful post-hoc explanations on top of the knowledge graphs. The expected contributions of this doctoral proposal are: 1) A novel framework to integrate machine learning methods and *KG*; 2) Explainability with the exploitation of the semantics of an entity in *KG*; 3) Exploitation of cause-effect relationships to empower the explanations; 4) Formalism for the metrics to evaluate the explanations and enhance the efficiency.

[4] Analogously Head prediction $\langle ?,\ r,\ e_t\ \rangle$.

4 Research Methodology

The research methodology for this Ph.D. proposal refers to a structured, conceptual, analysis of the Semantic Web Technologies applied to the research problem of explainability. The **RQ3)** is still an open problem, the first step focuses on the integration of ontology and entailment regimes, i.e., `RDFS and OWL`. For instance, extracting the explanations of tail prediction shown in Fig. 1, the baseline ignores the ontology in their approach. Considering this metadata of the entity, we aim to enhance the explainability of a model's prediction. The second step involves logical reasoning over the *KG*, considering the most implicit facts, and adding those facts to any other training entities would enhance the tail rank for that particular entity. This shows the impact of implicit knowledge in explainability (**RQ3)**. Our goal is not to limit the approach to RDFS entailment but also to extend it to OWL entailment in the next steps of the Ph.D.
The third step aims at integrating the machine learning frameworks with *KGs* to provide more insights into explainability (**RQ1)**. Three main components were identified for the implementation of explainable driven frameworks over *KG*: collecting valid data, training the model, and creating the explanations. The data collected from the RDF graph, given to any machine learning model needs to be valid, so our approach first implements the SHACL constraints to assure validity. One technique would be adapting SPARQL queries over an RDF graph, to avoid inconsistencies in the data given to the predictive model and the second technique will be reasoning via SPARQL queries to retrieve the implicit knowledge of a particular prediction. In the end, our approach will generate a knowledge graph comprising all the traced metadata collected during a data-driven pipeline followed during the resolution of a prediction task.
LIME [11] and Kelpie [13] can be the basis for providing feature contributions and fact-based explanations. **RQ2)** still remains a complex challenge. For instance, consider explainability task when integrated with *KGs* generates a huge search space for selecting a particular feature or a fact which leads the model to predict. Henceforth, the main aim of **RQ2)** is to provide valid scalable and trustable explanations. To tackle **RQ2)**, the implementation of an algorithm that optimally prunes the search space of valid explanations will be accomplished in the subsequent phases of this Ph.D. proposal. The fourth step includes the exploitation of causal relationships between the entities and provides more insights into the extraction of explanations. The last step integrates all the components and implements a fully-fledged framework compatible with domain agnostic.

5 Evaluation Plan

The previous section outlines the benefit of integrating knowledge-driven frameworks with explainable frameworks. Indeed **RQ1)** is a non-trivial research question, it might be complicated to evaluate a framework as a complete end-to-end task. However, the aim of **RQ1)** will be to integrate the symbolic system with sub-symbolic approaches. The creation of the possible set of explanations **RQ2)**

can be decomposed into two categories: SHACL validation result, then building the possible explanation set. To evaluate the effectiveness of explanations, metrics like `Mean Reciprocal Rank (MRR)` or `Hits@k` will be utilized.
Indeed, for the evaluation of predictive models, i.e., supervised learning, metrics like Accuracy, Precision, or Recall are computed. To explain predictive models, LIME [11] would be a better option to have an idea of influential features for a particular prediction. In the end, the aim will be to build a knowledge graph with all the characteristics traced in the predictive task (i.e., features, prediction probabilities, etc.) to enhance the explainability of the particular entity. The techniques to evaluate the framework with axioms injection **RQ3)**, would be challenging to provide new insights. **RQ3)** will also attempt to define and formalize the metrics quantifying the enhanced performance of explainability.
We aim to evaluate **InterpretME** on top of the following *KGs*: 1) *ImProVIT*[5] to explain the link prediction task about the impact of the immune system over the response of Hepatitis B and Influenza vaccines; 2) *CLARIFY*[6] to define machine learning models to predict biomarkers of a lung cancer patient and generate explanations. In these tasks, the components of InterpretME will be evaluated and implemented to provide explainability over *KGs*; assist the domain experts to have more insights into the predictive task.

6 Results So Far

We have analyzed the state of the art and the challenges to achieve explainability in prediction models. As a first result, we have developed **InterpretME**, a tool describes in a *KG* the outcomes and interpretations of a predictive method. In this section, we report on our initial assessment of **InterpretME** in the explainability of link prediction (Fig. 2). The *French Royalty KG* [7] depicts the information about each person in the French royal family.

LP model	Hits@1	MRR	Score	rank
Necessary	0.07	-0.02	-5.84	1
Sufficient	0.093	0.39	-7.47	3

(a) Link Prediction(LP) model performance

InterpretME	ΔHits@1	ΔMRR	Score	rank
Necessary	-0.023	-0.02	**-7.47**	**3**
Sufficient	0.10	0.0052	**-6.06**	**1**

(b) Evalution of InterpretME explanations

Fig. 2: **Initial Results**. Fig. 2a shows the evaluation of link prediction model performed over French Royalty KG. Fig. 2b indicates the effectiveness of removing or adding the combination of facts. The values in *bold* indicate the change in the metrics. For *necessary*, score and tail rank are worsened, and conversely for *sufficient* got improved. We report the efficacy of explanations as the difference of `Hits@1` and `MRR` on the particular tail prediction i.e., Δ`Hits@1` and Δ`MRR`. For necessary, more negative values $\Rightarrow$ higher efficacy and, sufficient, more positive values $\Rightarrow$ higher efficacy.

[5] German Funded project https://www.tib.eu/en/research-development/project-overview/project-summary/improvit
[6] EU H2020 Funded project https://www.clarify2020.eu/

Explainability over French Royalty KG[7] The link predictive task is to determine whether a member of the French royalty has a spouse. For instance, let us consider the tail prediction of a french royal member ⟨dbo:Charles the Simple, dbo:hasSpouse, dbo:Yes⟩ over *French Royalty KG*[8]. The evaluation of the *LP* model obtained respectively with `TransE` is reported. Here, the *necessary* explanations are analogous to the state-of-the-art approaches with the removal of combination facts. *Sufficient* explanations are given based on axioms injected, showing how to improve the predictions by adding implicit knowledge to any random entities in the training. Removing only the most important fact about the `dbo:Charles the Simple` will likely not change the prediction because it is still supported by other facts. Hence, **InterpetME** identifies the *necessary* explanations, removing the combination of facts featuring e_h, i.e., *Charles the Simple has child Gisela of France and has a spouse Frederuna* leads to worsens ϕ. The `score` reduced drastically to 7.47 and the `tail rank` from 1 to 3.
Sufficient scenario, **InterpretME** utilizes the implicit knowledge encoded in an ontology of the *KG* using entailment regimes, i.e., *subClassOf*, *domain*, *subPropertyOf*. For instance, using ontology about *dbo:Charles the Simple*, the implicit facts obtained: ⟨*dbo:marriedTo, rdfs:subPropertyOf, dbo:spouse*⟩, and ⟨*dbo:child, rdfs:domain,dbo:Person*⟩ added to the entities improve the *score* and the *tail rank* change from 3 to 1. For instance, assume the fact: ⟨dbo:Charles the Simple, dbo:marriedTo, dbo:Eadgifu of Wessex⟩, with the entailment regime *subPropertyOf*. As a result, we can infer: ⟨dbo:Charles the Simple, dbo:spouse, dbo:Eadgifu of Wessex⟩. Since the original model was not able to infer the correct tail, their `Hits@1` and `MRR` are more likely to be null. The re-trained should infer the correct predict tail, i.e., `dbo:Yes`. Thus, adding the implicit fact improves the prediction of an entity. The variation observed in the `Hits@1` and `MRR` are reported. As the *KG* is limited to some instances, the neighborhood of that particular member is less. Lastly, the SHACL validation results are used to provide one more layer of explainability and to justify the model's outcome. Here, SHACL constraint explains a `dbo:Person` link satisfying the domain constraints about a member having a child or spouse. For instance, the tail prediction ⟨`dbo:Charles the Simple, dbo:hasSpouse, dbo:Yes`⟩, the head entity satisfies the constraint having a child. The initial results are prominent to address the **RQ3)** and more refined results will be achieved in the next steps of this doctoral work.

7 Conclusions and Lessons Learned

This proposal introduces and formalizes the problem of explainability which can be particularly useful for explaining the Link Prediction model over *KGs*. In state-of-art approaches, we have explored different techniques for extracting explanations, including *LIME* and *Kelpie*. We found that each of these techniques has strengths and weaknesses depending on the application domain. One area for improvement is to prune the search space for explanations, and the techniques

[7] https://github.com/SDM-TIB/LinkPrediction_Explanations_over_KGs
[8] https://labs.tib.eu/sdm/InterpretME-og/sparql

for handling the SPARQL queries and entailment regimes. The aim was to see the impact of considering the axioms in the explainability of a Link Prediction model. The proposal identifies the challenges for developing a framework with exploiting semantics over the *KGs* to show more expressiveness in the explainability. We expect that the proposed research will make contributions to the development of a more robust explainable framework over *KGs*.
The next task for my Ph.D. will be to improvise the proposed approach for enhancing the explainability of the machine learning models over the *KGs*. The future work will be about the execution of a fully fledge explainable framework, like searching relevant entailment regimes or important characteristics of an entity. Moreover, the presented research plan involves the cause-effect relations between entities. Henceforth, future work will also be on such causal relations.

Acknowledgements I express my special thanks to my supervisor Prof. Dr. Maria-Esther Vidal her guidance and support. This work is funded by TrustKG-Transforming Data in Trustable Insights with grant P99/2020.

References

1. Akrami, F., Saeef, M.S., Zhang, Q., Hu, W., Li, C.: Realistic re-evaluation of knowledge graph completion methods: An experimental study. In: SIGMOD (2020)
2. Beckers, S.: Causal explanations and XAI. CLeaR (2022)
3. Borrego, A., Ayala, D., Hernández, I., Rivero, C.R., Ruiz, D.: Cafe: Knowledge graph completion using neighborhood-aware features. IFAC (2021)
4. De Bie, T., De Raedt, L., Hern_'andez-Orallo, J., Hoos, H.H., Smyth, P., Williams, C.K.I.: Automating data science. CACM (2022)
5. Figuera, M., Rohde, P.D., Vidal, M.E.: Trav-shacl: Efficiently validating networks of shacl constraints. In: WWW (2021)
6. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D.: A survey of methods for explaining black box models. ACM Comput. Surv. (2018)
7. Halliwell, N., Gandon, F., Lecue, F.: User Scored Evaluation of Non-Unique Explanations for Relational Graph Convolutional Network Link Prediction on Knowledge Graphs. In: K-CAP. ACM (2021)
8. Huang, H.: Causal relationship over knowledge graphs. In: CIKM (2022)
9. Myszczynska, M., Ojamies, P., Lacoste, A., Neil, D., Saffari, A., Mead, R., Hautbergue, G., Holbrook, J., Ferraiuolo, L.: Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nature Reviews Neurology (2020)
10. Pearl, J.: Theoretical impediments to machine learning with seven sparks from the causal revolution. WSDM (2018)
11. Ribeiro, M.T., Singh, S., Guestrin, C.: "why should i trust you?": Explaining the predictions of any classifier. In: SIGKDD (2016)
12. Rohde, P.D.: SHACL Constraint Validation during SPARQL Query Processing. In: VLDB PhD Workshop (2021)
13. Rossi, A., Firmani, D., Merialdo, P., Teofili, T.: Explaining link prediction systems based on knowledge graph embeddings. In: SIGMOD (2022)
14. Salimi, B., Parikh, H., Kayali, M., Getoor, L., Roy, S., Suciu, D.: Causal relational learning. In: SIGMOD (2020)
15. Zhang, H., Zheng, T., Gao, J., Miao, C., Su, L., Li, Y., Ren, K.: Data poisoning attack against knowledge graph embedding. In: IJCAI (2019)