HHAI 2023: Augmenting Human Intellect
P. Lukowicz et al. (Eds.)

doi:10.3233/FAIA230091

Towards Explainable Automatic Knowledge Graph Construction with Human-in-the-Loop

Bohui ZHANG [a,1], Albert MEROÑO PEÑUELA [a] and Elena SIMPERL [a]
[a] *Department of Informatics, King's College London*
ORCiD ID: Bohui Zhang https://orcid.org/0000-0001-5430-1624, Albert Meroño Peñuela https://orcid.org/0000-0003-4646-5842, Elena Simperl https://orcid.org/0000-0003-1722-947X

Abstract. Knowledge graphs are important in human-centered AI because of their ability to reduce the need for large labelled machine-learning datasets, facilitate transfer learning, and generate explanations. However, knowledge-graph construction has evolved into a complex, semi-automatic process that increasingly relies on opaque deep-learning models and vast collections of heterogeneous data sources to scale. The knowledge-graph lifecycle is not transparent, accountability is limited, and there are no accounts of, or indeed methods to determine, how fair a knowledge graph is in the downstream applications that use it. Knowledge graphs are thus at odds with AI regulation, for instance the EU's upcoming AI Act, and with ongoing efforts elsewhere in AI to audit and debias data and algorithms. This paper reports on work in progress towards designing explainable (XAI) knowledge-graph construction pipelines with human-in-the-loop and discusses research topics in this space. These were grounded in a systematic literature review, in which we studied tasks in knowledge-graph construction that are often automated, as well as common methods to explain how they work and their outcomes. We identified three directions for future research: (i) tasks in knowledge-graph construction where manual input remains essential and where there may be opportunities for AI assistance; (ii) integrating XAI methods into established knowledge-engineering practices to improve stakeholder experience; as well as (iii) evaluating how effective explanations genuinely are in making knowledge-graph construction more trustworthy.

Keywords. knowledge graph, knowledge-graph construction, knowledge engineering, transparency, explainability, XAI

1. Introduction: Raising Concerns of Knowledge-Graph Transparency

To reach its potential, AI needs data and context. Without the right (amounts of) data, machine learning (ML) cannot identify patterns or make predictions. Without a deeper understanding of context, AI applications cannot engage people in a meaningful way. Knowledge graphs (KGs) [37], a term coined by Google in 2012 to refer to its general-purpose knowledge base, are critical to both: they reduce the need for large labelled ML

[1]Corresponding Author: Bohui Zhang, Email: bohui.zhang@kcl.ac.uk.

datasets, facilitate transfer learning, and generate explanations [106]. KGs are routinely used alongside ML in many applications, including search, question answering, recommendation [37] and, in industry contexts, enterprise data management, digital twins, supply chain management, procurement, and regulatory compliance [89].

As AI applications produce and consume more data, engineering KGs has evolved into a complex, semi-automatic process that increasingly relies on opaque deep-learning models and vast collections of heterogeneous data sources to scale to graphs with millions of entities and billions of statements [104,119]. The KG lifecycle is not transparent [121], accountability is limited, and accounts of how biased a KG is [1] or how fair in the downstream applications that use it [29] are patchy. KGs are thus at odds with AI regulation, for instance the EU's upcoming AI Act,[2] and with ongoing efforts elsewhere in AI to systematically audit and debias data and algorithms [11,19,38,71,78].

Regulators take a risk-based approach to the use of AI, prescribing, among other things, transparency and accountability obligations for different classes of AI applications. Organisations using KGs, either directly as data infrastructure, or as graph embeddings in ML models, will face challenges unless they can document and attest that their KGs are compliant with the law. Furthermore, when a KG is part of an AI application that counts as high-risk, that application will have to undergo conformity assessments both at design and at run time. KGs themselves are meant to make ML models explainable [106] and hence facilitate such compliance tasks, but that would imply that the KG lifecycle abides by the same rules.

We argue that this is not yet the case. With this paper, we would like to advance the vision of **trustworthy KG engineering** to allow KG stakeholders to rely appropriately on AI algorithms and use KGs with confidence [50]. For this to happen, we need to first gain a better understanding of emerging knowledge-graph construction (KG construction) practices in the era of ML-as-a-service and develop human-in-the-loop approaches to ensure transparency and accountability throughout the KG lifecycle. This applies to both proprietary KGs used within organisations [89] and publicly available KGs like Wikidata [110], DBPedia [5], YAGO [102], ConceptNet [98], which are extensively used by researchers and practitioners. As AI laws and regulations enter into force, the trustworthy credentials of such KGs will have to be systematically assessed and documented.

Our paper follows from recent work that explored emergent neuro-symbolic AI architectures from a system-design perspective. Van Bekkum et al. [109] proposed a taxonomy of hybrid (i.e., learning and reasoning) systems and discussed common architecture patterns and use cases. Building on their insights, Breit et al. [14] carried out a comprehensive literature review to add details to those patterns in terms of inputs, outputs, processing units, types of ML models and their training, types of knowledge representation and reasoning, but also transparency and auditability. One of their main findings was that most system designers do not consider these latter aspects at all, or, when they do, that they do not evaluate them sufficiently. A third paper by Tamašauskaitė and Groth [104] drew from a survey of system papers to define a canonical KG construction process. Our work continues where they left off: starting from their KG construction process, we follow one of their main recommendations to map tools and techniques for each step to provide additional guidance to researchers and developers. We analyse the KG lifecycle to identify tasks are are commonly automated with AI and those which still require

[2] https://artificialintelligenceact.eu/

human input and oversight and could potentially benefit from AI assistance. In parallel, we survey the state of the art in explainable AI (XAI) to inform the design of XAI approaches that are genuinely useful for KG stakeholders such as knowledge engineers, subject domain experts, and users. Our main findings are:

1. There are tasks in KG construction, for instance knowledge acquisition, where automation[3] is routinely used with promising results. At the same time, there are opportunities to use AI to assist other tasks such as ontology reuse, ontology evolution, ontology evaluation, documentation etc, where (the latest) AI capabilities have remained under-explored.
2. While tasks around knowledge acquisition, taxonomy building, and data ingestion are often automated, human oversight is still needed to improve performance, establish trust, or comply with the law. In our review we found little evidence of integration of AI capabilities, no matter their level of interpretability, into standard knowledge-engineering tools and practices. Furthermore, our understanding of human-in-the-loop KG construction remains limited, with implications for user experience.
3. Comprehensive evaluations of XAI methods are lacking, with most studies focusing on simple ML models in lab settings, with mixed results [73,97,116]. The KG community, just like elsewhere in AI, needs to gain a better understanding of how people react to and use explanations to build trust and boost technology adoption.

Based on these findings we propose several directions for future research, drawing on theory and insights from AI, but also human-AI interaction [3], interactive ML [28], and social computing [84,93]. These include: (i) AI assistants for overlooked tasks in the KG lifecycle; (ii) end-to-end tools supporting automated KG construction with human-in-the-loop with built-in advanced, explainable AI capabilities; as well as (iii) holistic evaluation frameworks that assess the extent to which explanations genuinely help humans engineer better KGs.

2. Background: Knowledge Engineering, Knowledge Graphs, and Transparency

Knowledge engineering, the branch of AI concerned with building and managing knowledge-based systems [87,100], has changed dramatically with the latest innovations in machine learning, natural language processing, and computer vision. And yet, as the most recent advances in large language models and generative AI demonstrate, the question of how to capture and encode domain knowledge into a computational representation remains as challenging as ever [85]. The technologies and end-user tools to support core knowledge-engineering tasks such as knowledge acquisition have advanced significantly to meet the scale requirements of modern KGs [131]. At the same time, the most effective approaches to knowledge representation still require human oversight at various levels [94,95], but increasingly human input is in the form of augmenting or validating algorithmic suggestions [104].

[3]In this paper we use AI assistance and automation interchangeably. While we acknowledge that not all automation in KG construction is AI, we argue that the use of AI brings about specific challenges with respect to transparency, accountability etc.

Knowledge graphs are just one of the latest manifestation of knowledge engineering, alongside property graphs [4], and before them ontologies [44] and knowledge bases [46]. They use a schema or ontology to organise data and reason over it to infer new facts and flag inconsistencies [37]. While there are various knowledge-graph definitions, most agree on the following attributes, which distinguish them from technologies like relational databases and semantic networks: first, data is organised in a directed, labelled graph. Nodes are entities of interest in a domain and their abstract classes. Edges stand for relationships and attributes between them. Like classes, relationships and attributes can be arranged in a taxonomy. They can also have features like transitivity, domain or range restrictions, etc. Second, graph labels have well-defined meanings for programmatic use in data validation and reasoning. Nodes and edges are accessed through unique identifiers such as web URIs. Many KG representational languages exist, each with its own formal semantics and syntax (e.g., W3C RDF[4], RDFS[5]). Finally, KGs are general-purpose knowledge bases, meant to be used by multiple applications. They evolve to accommodate changes in the domain, data, and user requirements. In KG engineering, it is best practice to reuse external ontologies and define links from one KG to another to speed up development and facilitate data interoperability.
Transparency as an AI design principle stands for the need to clearly document and explain how an AI system makes decisions, how the data is collected, used, and governed, and how the system is evaluated and audited [27,42,47]. One of the key mechanisms to achieve transparency is explainability of ML models. While some ML models such as decision trees could be considered interpretable by design, others such as large language models are too complex for people to comprehend in the same way. Within the context of trustworthy AI, researchers and practitioners have proposed many XAI frameworks, guidance, standards [88], techniques [57,77], and evaluation metrics [35] for various models. Among them, some suggested to use KGs to generate explanations. By exploring paths in KG and formatting them into natural language justification, Silva et al. [92] inject interpretability into text entailment system. Another example is for computer vision, where Wang et al. [115] distill information from both word embeddings and knowledge-graph representations for zero-shot recognition.

3. The KG Lifecycle

Building on the process from [104], Figure 1 shows an exemplary KG construction pipeline with a mix of automated and manual capabilities and contributions from several stakeholder groups: knowledge engineering and machine learning specialists, subject domain experts, online volunteers and crowdsourcing services, as well as developers of applications using KGs.

As the figure suggests, **KGs are interacting with AI capabilities in complex ways**. On the left (1), multiple data sources, structured and unstructured, are lifted into KGs using ML for named entity recognition [128], relation extraction [54], entity reconciliation [90], link prediction [80] and many others. The ontology organising the KG can be provided upfront or derived from the data itself, depending on whether there is a clear domain or available structured data with predefined types of entities and relations [104].

[4] https://www.w3.org/RDF/
[5] https://www.w3.org/TR/rdf-schema/

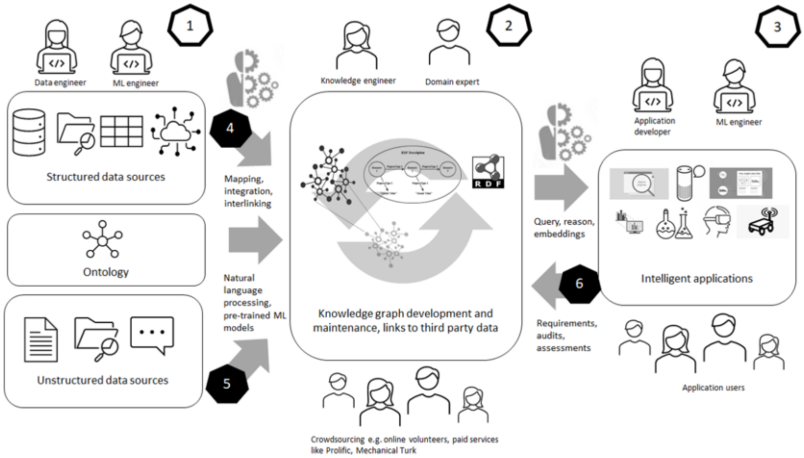

Figure 1. The KG lifecycle.

In this context, [121] discusses the need for more transparency with respect to data provenance and currency; both can affect whether application developers will be able to use the KG with confidence as a source of reliable, complete, unbiased, up-to-date information. The result of knowledge acquisition is shown in the middle of the figure (2), where KGs are often linked to third-party data, reuse standard ontologies and identifiers, and are encoded as RDF, JSON or other formats. On the right-hand side of the figure (3), there is a selection of use cases for KGs alongside other forms of AI. KGs are used as knowledge bases to query and reason upon, for instance in search [120], question answering [16,32], or recommendation [140]. Information can be obtained from a graph through deductive (e.g., logical rules) and inductive methods (e.g., as continuous graph embeddings) [37]. Both methods need to be transparent to the user [13,81] to be trustworthy.

KG maintenance is prompted by source updates on the left (1), and requirements, audits, and assessments on the right (3). Human-in-the-loop tasks (arrows in the figure $(4-6)$) increasingly use ML models with varying levels of interpretability. Crowdsourcing for supervising ML (bottom-middle of the figure) has similar transparency challenges as the algorithms it complements. This is because the digital services commonly used for this purpose e.g., Prolific, Mechanical Turk, are black-box, proprietary platforms with limited means to replicate or reproduce results [74].

4. XAI in the KG Lifecycle

Following the discussion of the lifecycle, we carried out a PRISMA [70] literature review on databases including ACM Digital Library, IEEExplore, ScienceDirect, arXiv, SpringerLink, and Google Scholar. We searched for queries combining, on the one side, keywords related to transparency (transparent, transparency, interpretable, interpretability, explainable, explainability) and, on the other side, keywords related to KG construction (knowledge graph construct*, knowledge graph develop*, knowledge graph complet*, knowledge graph refine*, knowledge graph reasoning, knowledge graph inference, knowledge engineering) and tasks (named entity recognition, extract entities, relation extraction, entity linking, entity matching, entity resolution, entity alignment, link prediction). The search took place from October to December 2022 and resulted in more than 735 thousand hits. We then took the top 50 hits per query, which led to around four thousand papers, with duplicates.[6] We assessed relevance based on titles, abstracts, and keywords first, and in a second step, reviewed the text of the paper to select only those papers which proposed a solution to transparent KG construction, either as a whole process or for individual tasks. We discarded papers that only mentioned transparency and related concepts rather than putting forward a solution. The final corpus consisted of 84 papers. The papers were all published in the past ten years, which was to be expected given the term ”knowledge graphs” was coined in 2012 and is inline with other recent knowledge-graph surveys [86,104].

The authors classified the papers reviewed with respect to KG construction tasks they addressed and their approach to explainability, starting with categories widely used in the literature. For explainability we started with what is explained: local (data point) vs global (outcome) and when: post-hoc (after prediction) vs self-explaining (while predicting, or inherently interpretable). For post-hoc models, another layer of coding is added for both local and global explanation methods to consider whether the XAI methods are independent of the ML models or not: model-agnostic (can be applied to any ML models) vs model-specific (explicitly designed for a specific (group of) model architecture(s)). Finally, because we also checked the extent to which the solutions considered human-AI interaction aspects, for instance by proposing specific affordances for people to engage with the explanations in some way, as opposed to the explanation being merely communicated to (an unspecified group of) users.

The result of the classification is presented in Table 1. At a glance, the papers we reviewed do not cover the entire KG lifecycle. Most papers are concerned with knowledge acquisition via entity extraction (as a source of classes and instances in KGs) and relation extraction (as a source of property classes, but more importantly connecting entities to each other through properties), or with curation and maintenance via entity resolution (consolidating the data that refers to the same entities) and link prediction (suggesting missing or emerging facts). Besides the four tasks at the top of the table, we found one paper dealing with the evolution of the KG schema or ontology [61] and another one about detecting and explaining inconsistency in KGs [107]. We note that link prediction was by far the most popular task, and that a majority of papers dealt with curation and maintenance rather than building a KG for a particular purpose. This is somewhat concerning, as many applications of KGs are in enterprise contexts, where the first step is to build a computational representation of the enterprise's data, which is stored across various systems and modalities.

A second high-level observation is the balanced split in the chosen format for the explanations. Methods based on input and generated features use attention weights [41,141], words [49,52], attributes [8] etc. to generate explanations, which can be numerical, textual, or visual. By contrast, methods based on human-understandable background knowledge provide rules, reasoning paths, and structured contextual information as explanations. Given that we're interested in explanations that are accessible to knowledge engineers and subject domain experts, it would be interesting to evaluate if their familiarity with knowledge representation and/or the subject domain impacts how use-

[6]The six platforms where we performed the search supported different query affordances. This means that in some cases it was possible to build complex queries with multiple keyword options, whereas in others we had to use separate queries to achieve the same results. We took the first 50 hits for each search query.

		Entity Extraction	Relation Extraction	Entity Resolution	Link Prediction	Others
Local post-hoc	Model-agnostic			LEMON [8], Minun [113], Landmark [7], CERTA [105], Mojito [22]		
	Model-specific			LightEA [59], HIF+KAT [130], GMKSLEM [21]	XTransE [137], CPM [99], Kelpie [81], CrossE [138], KGInfluence [143], GINN [40], SNS [39], approxSemanticCrossE [18]	
Local self-explaining		TMN [52], BTPK [15], AutoTriggER [49], Instance-based [68]	[91], D-REX [2], SIRE [134], SAIS [125], NERO [141], DISCO-RE [111], SemRep [45], LogiRE [82]	XINA [132], RuleSynth [96]	GCNN w/ att [62], T-GAP [41], TAGAT [117], DisenKGAT [123], IDEAL [133], ITCN [122] [12], [136], AnyBURL [60], xERTE [34], DRUM [83], TITer [103], SAFRAN [67], MINERVA [20], TLogic [55], DeepPath [127], SLiCE [114], PathCon [112], CogKR [25], Gradient Rollback [48], RNNLogic [76], CPL [30], GPFL [31], CAKE [64], R2D2 [36], RED-GNN [139], HiAM [58], SparKGR [124], [65], [142], RuleGuider [51], RPJE [66], RuleDict [135], [10], NTPs [79], SQUIRE [6], LCGE [63]	
Global post-hoc	Model-agnostic			ExplainER [26]		
	Model-specific	Emboot [144]	ProtoRE [24]		MGNN [17], CRIAGE [72], ITransF [126], DensE [56], HopfE [9], METransE [118]	
Global self-explaining				xER [108]	FTL-LM [53], Neural LP [129]	[61], Abstraction [107]
Human-in-the-loop		[43]		SystemER [75], TuneR [69]		

Table 1. Overview of explainable knowledge graph construction methods. We add an additional class for human-in-the-loop methods except for the four main categories.

ful knowledge-based explanations are compared to feature-based ones, which sometimes require an understanding of machine learning. At the same time, explanations are generated in a different way for each of the four core KG construction tasks at the top of the table. For entity and relation extraction, explanations often refer to contextual cues such as triggers [49,52] and sentences [91]. Explanations for entity resolution tend to use entity matching rules [69,75] and (ranked) attributes of the entity pair [8,26]. Finally, link prediction methods use the topology and reasoning capabilities of the KG. Rule- and path-based methods have become the majority format of explanations, achieved through random walk-based methods [53,55], reinforcement learning [36,60,127], and perturbation-based methods [72], etc.

There are very few papers considering human inputs or oversight, which are critical in trustworthy AI frameworks and guidance [23]. Human input in isolated cases [43,69,75] often involves providing or revising rules for tasks like entity resolution. Furthermore, most approaches have not been comprehensively evaluated. The majority of methods (58 out of 84) do not perform evaluation or use informal evaluations by visualizing and commenting on a limited number of cases of explaining outcomes intuitively. Only a few of them include user study (or human evaluation) and task-specific metrics.

5. Directions for Future Research

There are three directions for future research that follow from the review. First, going back to prior literature on knowledge engineering methodologies [44,87,100,101], there are many tasks and activities where automation remains an exception. Aside from the four tasks at the top of Table 1, there is an opportunity to think about other ways for AI assistance to add value: for instance, one design principle of KGs is that they are meant to integrate across multiple sources and be able to tackle evolving requirements. Reusing existing schemas or ontologies can help with interoperability, but the task of finding or assessing an ontology for reuse is still mostly manual. At the other end of the lifecycle, documenting KGs can help with maintenance and reuse, and advances in generative AI make it a chief candidate for automation. While we found a range of explainable link prediction approaches, it would be useful to dive deeper into this sub-field to understand the extent to which these different approaches solve common concerns around the quality of KGs. One difference between representing knowledge in a KG and a machine-learning model is that a KG can provide guarantees about the validity of the information, its provenance, its currency, etc. upon retrieval. However, this is predicated by KGs being regularly audited according these and other quality dimensions and improved. Link prediction is one way to do this, alongside many others, e.g., debiasing [29]. Furthermore, while knowledge acquisition is generally well represented in the literature, a lot of work focuses on text rather than other data modalities, which is a concern in many KG application areas, e.g., enterprise data management (which needs to work with structured data) or cultural heritage (where a lot of domain data is neither text nor numbers).

Second, as we noted earlier, the fewest of approaches look at the human-in-the-loop aspects of KG construction, including human agency and oversight, feedback, etc [23]. While there is a lot of work in human-AI interaction and interactive ML in the HCI community, they tend to focus so far on simpler ML models and different applications that the knowledge production scenarios we are interested in. One exception is the work on ORES [33], a participatory ML system used in Wikipedia and Wikidata (a large open-source KG). However, the Wikidata KG construction process is quite unique because it is community-based, with more than 24 thousand active contributors[7] who receive AI assistance for distinct tasks such as vandalism detection and consistency checks. We need to follow their example to develop the same types of workflows and tools for other KG construction scenarios - in most cases, these involve much smaller teams and different tool environments. The majority of existing integrated development environments (IDE) for KGs (e.g. PoolParty[8], data.world[9], Protégé[10]) assume KGs are mostly built manually, with some basic automation to speed-up routine tasks like translating node labels or creating documentation from node and edges descriptions. Large language models like ChatGPT offer chances to develop novel KG editing tools and interactions, allowing people to interact with their AI assistants via natural language and ensuring transparency. Meanwhile, developers working with KGs require KG-related process blueprints that utilize AI algorithms and adhere to AI regulations for creating downstream applications.

Thirdly, our review flagged the need for better evaluations, which encompasses metrics, benchmarks, and datasets, as well as toolkits and guidance for conducting studies that assess how effective the explanations supplied in KG construction tasks are as proxies and enablers for transparent and hence trusted KGs.

[7] https://www.wikidata.org/wiki/Wikidata:Statistics
[8] https://www.poolparty.biz/
[9] https://data.world/
[10] https://protege.stanford.edu/

References

[1] David Abián, Albert Meroño-Peñuela, and Elena Simperl. An analysis of content gaps versus user needs in the wikidata knowledge graph. In *The Semantic Web–ISWC 2022: 21st International Semantic Web Conference, Virtual Event, October 23–27, 2022, Proceedings*, pages 354–374. Springer, 2022.

[2] Alon Albalak, Varun Embar, Yi-Lin Tuan, Lise Getoor, and William Yang Wang. D-REX: dialogue relation extraction with explanations. *CoRR*, abs/2109.05126, 2021.

[3] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. Guidelines for human-ai interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*, pages 1–13, 2019.

[4] Renzo Angles. The property graph database model. In *AMW*, 2018.

[5] Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. Dbpedia: A nucleus for a web of open data. In *Proceedings of the 6th International The Semantic Web and 2nd Asian Conference on Asian Semantic Web Conference*, ISWC'07/ASWC'07, page 722–735, Berlin, Heidelberg, 2007. Springer-Verlag.

[6] Yushi Bai, Xin Lv, Juanzi Li, Lei Hou, Yincen Qu, Zelin Dai, and Feiyu Xiong. SQUIRE: A sequence-to-sequence framework for multi-hop knowledge graph reasoning. *CoRR*, abs/2201.06206, 2022.

[7] Andrea Baraldi, Francesco Del Buono, Matteo Paganelli, and Francesco Guerra. Landmark explanation: An explainer for entity matching models. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, CIKM '21, page 4680–4684, New York, NY, USA, 2021. Association for Computing Machinery.

[8] Nils Barlaug. LEMON: explainable entity matching. *CoRR*, abs/2110.00516, 2021.

[9] Anson Bastos, Kuldeep Singh, Abhishek Nadgeri, Saeedeh Shekarpour, Isaiah Onando Mulang, and Johannes Hoffart. Hopfe: Knowledge graph representation learning using inverse hopf fibrations. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, CIKM '21, page 89–99, New York, NY, USA, 2021. Association for Computing Machinery.

[10] Patrick Betz, Christian Meilicke, and Heiner Stuckenschmidt. Supervised knowledge aggregation for knowledge graph completion. In Paul Groth, Maria-Esther Vidal, Fabian Suchanek, Pedro Szekley, Pavan Kapanipathi, Catia Pesquita, Hala Skaf-Molli, and Minna Tamper, editors, *The Semantic Web*, pages 74–92, Cham, 2022. Springer International Publishing.

[11] Lucas Beyer, Olivier J Hénaff, Alexander Kolesnikov, Xiaohua Zhai, and Aäron van den Oord. Are we done with imagenet? *arXiv preprint arXiv:2006.07159*, 2020.

[12] Rajarshi Bhowmik and Gerard de Melo. A joint framework for inductive representation learning and explainable reasoning in knowledge graphs. *CoRR*, abs/2005.00637, 2020.

[13] Federico Bianchi, Gaetano Rossiello, Luca Costabello, Matteo Palmonari, and Pasquale Minervini. Knowledge graph embeddings and explainable AI. *CoRR*, abs/2004.14843, 2020.

[14] Anna Breit, Laura Waltersdorfer, Fajar J. Ekaputra, Marta Sabou, Andreas Ekelhart, Andreea Iana, Heiko Paulheim, Jan Portisch, Artem Revenko, Annette ten Teije, and Frank van Harmelen. Combining machine learning and semantic web: A systematic mapping study. *ACM Comput. Surv.*, mar 2023. Just Accepted.

[15] Yulin Chen, Zelai Yao, Haixiao Chi, Dov M. Gabbay, Bo Yuan, Bruno Bentzen, and Beishui Liao. Btpk-based learning: An interpretable method for named entity recognition. *CoRR*, abs/2201.09523, 2022.

[16] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Ben Hutchinson, Reiner Pope, James Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier Garcia, Vedant Misra, Kevin Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito, David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayana Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Diaz, Orhan Firat, Michele Catasta, Jason Wei, Kathy Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. Palm: Scaling language modeling with pathways, 2022.

[17] David Jaime Tena Cucala, Bernardo Cuenca Grau, Egor V. Kostylev, and Boris Motik. Explainable GNN-based models over knowledge graphs. In *International Conference on Learning Representations*,

2022.

[18] Claudia d'Amato, Pierpaolo Masella, and Nicola Fanizzi. An approach based on semantic similarity to explaining link predictions on knowledge graphs. In *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, WI-IAT '21, page 170–177, New York, NY, USA, 2022. Association for Computing Machinery.

[19] David Danks and Alex John London. Algorithmic bias in autonomous systems. In *Ijcai*, volume 17, pages 4691–4697, 2017.

[20] Rajarshi Das, Shehzaad Dhuliawala, Manzil Zaheer, Luke Vilnis, Ishan Durugkar, Akshay Krishnamurthy, Alexander J. Smola, and Andrew McCallum. Go for a walk and arrive at the answer: Reasoning over paths in knowledge bases using reinforcement learning. *CoRR*, abs/1711.05851, 2017.

[21] Ting Deng, Lei Hou, and Ziyan Han. Keys as features for graph entity matching. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, pages 1974–1977, 2020.

[22] Vincenzo Di Cicco, Donatella Firmani, Nick Koudas, Paolo Merialdo, and Divesh Srivastava. Interpreting deep learning models for entity resolution: An experience report using lime. aiDM '19, New York, NY, USA, 2019. Association for Computing Machinery.

[23] Virginia Dignum. *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer Publishing Company, Incorporated, 1st edition, 2019.

[24] Ning Ding, Xiaobin Wang, Yao Fu, Guangwei Xu, Rui Wang, Pengjun Xie, Ying Shen, Fei Huang, Hai-Tao Zheng, and Rui Zhang. Prototypical representation learning for relation extraction. *CoRR*, abs/2103.11647, 2021.

[25] Zhengxiao Du, Chang Zhou, Jiangchao Yao, Teng Tu, Letian Cheng, Hongxia Yang, Jingren Zhou, and Jie Tang. Cogkr: Cognitive graph for multi-hop knowledge reasoning. *IEEE Transactions on Knowledge and Data Engineering*, 35(2):1283–1295, 2023.

[26] Amr Ebaid, Saravanan Thirumuruganathan, Walid G. Aref, Ahmed Elmagarmid, and Mourad Ouzzani. Explainer: Entity resolution explanations. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, pages 2000–2003, 2019.

[27] Upol Ehsan, Q Vera Liao, Michael Muller, Mark O Riedl, and Justin D Weisz. Expanding explainability: Towards social transparency in ai systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2021.

[28] Jerry Alan Fails and Dan R Olsen Jr. Interactive machine learning. In *Proceedings of the 8th international conference on Intelligent user interfaces*, pages 39–45, 2003.

[29] Joseph Fisher, Arpit Mittal, Dave Palfrey, and Christos Christodoulopoulos. Debiasing knowledge graph embeddings. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7332–7345, 2020.

[30] Cong Fu, Tong Chen, Meng Qu, Woojeong Jin, and Xiang Ren. Collaborative policy learning for open knowledge graph reasoning. *CoRR*, abs/1909.00230, 2019.

[31] Yulong Gu, Yu Guan, and Paolo Missier. Efficient rule learning with template saturation for knowledge graph completion. *CoRR*, abs/2003.06071, 2020.

[32] Jiaxian Guo, Junnan Li, Dongxu Li, Anthony Meng Huat Tiong, Boyang Li, Dacheng Tao, and Steven C. H. Hoi. From images to textual prompts: Zero-shot vqa with frozen large language models, 2022.

[33] Aaron Halfaker and R Stuart Geiger. Ores: Lowering barriers with participatory machine learning in wikipedia. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–37, 2020.

[34] Zhen Han, Peng Chen, Yunpu Ma, and Volker Tresp. xerte: Explainable reasoning on temporal knowledge graphs for forecasting future links. *CoRR*, abs/2012.15537, 2020.

[35] Peter Hase and Mohit Bansal. Evaluating explainable AI: Which algorithmic explanations help users predict model behavior? In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5540–5552, Online, July 2020. Association for Computational Linguistics.

[36] Marcel Hildebrandt, Jorge Andres Quintero Serna, Yunpu Ma, Martin Ringsquandl, Mitchell Joblin, and Volker Tresp. Reasoning on knowledge graphs with debate dynamics. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(04):4123–4131, Apr. 2020.

[37] Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, José Emilio Labra Gayo, Sabrina Kirrane, Sebastian Neumaier, Axel Polleres, Roberto Navigli, Axel-Cyrille Ngonga Ngomo, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan F. Sequeda, Steffen Staab, and Antoine Zimmermann. Knowledge graphs. *CoRR*, abs/2003.02320, 2020.

[38] Sara Hooker. Moving beyond "algorithmic bias is a data problem". *Patterns*, 2(4):100241, 2021.

[39] Md Kamrul Islam, Sabeur Aridhi, and Malika Smail-Tabbone. Negative sampling and rule mining for

explainable link prediction in knowledge graphs. *Knowledge-Based Systems*, 250:109083, 2022.

[40] Weihao Jiang, Yao Fu, Hong Zhao, Junhong Wan, and Shiliang Pu. Graph intention neural network for knowledge graph reasoning. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2022.

[41] Jaehun Jung, Jinhong Jung, and U Kang. Learning to walk across time for interpretable temporal knowledge graph completion. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, KDD '21, page 786–795, New York, NY, USA, 2021. Association for Computing Machinery.

[42] Davinder Kaur, Suleyman Uslu, Kaley J. Rittichier, and Arjan Durresi. Trustworthy artificial intelligence: A review. *ACM Comput. Surv.*, 55(2), jan 2022.

[43] Mayank Kejriwal, Runqi Shao, and Pedro Szekely. Expert-guided entity extraction using expressive rules. SIGIR'19, New York, NY, USA, 2019. Association for Computing Machinery.

[44] Elisa F Kendall and Deborah L McGuinness. Ontology engineering. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 9(1):i–102, 2019.

[45] Halil Kilicoglu, Graciela Rosemblat, Marcelo Fiszman, and Dongwook Shin. Broad-coverage biomedical relation extraction with SemRep. *BMC Bioinformatics*, 21(1):188, May 2020.

[46] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew Tomkins. Extracting large-scale knowledge bases from the web. In *VLDB*, volume 99, pages 639–650. Citeseer, 1999.

[47] Stefan Larsson and Fredrik Heintz. Transparency in artificial intelligence. *Internet Policy Review*, 9(2), 2020.

[48] Carolin Lawrence, Timo Sztyler, and Mathias Niepert. Explaining neural matrix factorization with gradient rollback. *CoRR*, abs/2010.05516, 2020.

[49] Dong-Ho Lee, Ravi Kiran Selvam, Sheikh Muhammad Sarwar, Bill Yuchen Lin, Mahak Agarwal, Fred Morstatter, Jay Pujara, Elizabeth Boschee, James Allan, and Xiang Ren. Autotrigger: Named entity recognition with auxiliary trigger extraction. *CoRR*, abs/2109.04726, 2021.

[50] John D Lee and Katrina A See. Trust in automation: Designing for appropriate reliance. *Human factors*, 46(1):50–80, 2004.

[51] Deren Lei, Gangrong Jiang, Xiaotao Gu, Kexuan Sun, Yuning Mao, and Xiang Ren. Learning collaborative agents with rule guidance for knowledge graph reasoning. *CoRR*, abs/2005.00571, 2020.

[52] Bill Yuchen Lin, Dong-Ho Lee, Ming Shen, Ryan Moreno, Xiao Huang, Prashant Shiralkar, and Xiang Ren. Triggerner: Learning with entity triggers as explanations for named entity recognition. *CoRR*, abs/2004.07493, 2020.

[53] Qika Lin, Rui Mao, Jun Liu, Fangzhi Xu, and Erik Cambria. Fusing topology contexts and logical rules in language models for knowledge graph completion. *Information Fusion*, 90:253–264, 2023.

[54] Yankai Lin, Shiqi Shen, Zhiyuan Liu, Huanbo Luan, and Maosong Sun. Neural relation extraction with selective attention over instances. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2124–2133, Berlin, Germany, August 2016. Association for Computational Linguistics.

[55] Yushan Liu, Yunpu Ma, Marcel Hildebrandt, Mitchell Joblin, and Volker Tresp. Tlogic: Temporal logical rules for explainable link forecasting on temporal knowledge graphs. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(4):4120–4127, Jun. 2022.

[56] Haonan Lu, Hailin Hu, and Xiaodong Lin. Dense: An enhanced non-commutative representation for knowledge graph embedding with adaptive semantic hierarchy. *Neurocomputing*, 476:115–125, 2022.

[57] Scott M. Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *CoRR*, abs/1705.07874, 2017.

[58] Ting Ma, Shangwen Lv, Longtao Huang, and Songlin Hu. Hiam: A hierarchical attention based model for knowledge graph multi-hop reasoning. *Neural Networks*, 143:261–270, 2021.

[59] Xin Mao, Wenting Wang, Yuanbin Wu, and Man Lan. Lightea: A scalable, robust, and interpretable entity alignment framework via three-view label propagation, 2022.

[60] Christian Meilicke, Melisachew Wudage Chekol, Manuel Fink, and Heiner Stuckenschmidt. Reinforced anytime bottom up rule learning for knowledge graph completion. *CoRR*, abs/2004.04412, 2020.

[61] Albert Meroño Peñuela, Romana Pernisch, Christophe Guéret, and Stefan Schlobach. Multi-domain and explainable prediction of changes in web vocabularies. In *Proceedings of the 11th on Knowledge Capture Conference*, K-CAP '21, page 193–200, New York, NY, USA, 2021. Association for Computing Machinery.

[62] Daniel Neil, Joss Briody, Alix Lacoste, Aaron Sim, Páidí Creed, and Amir Saffari. Interpretable graph convolutional neural networks for inference on noisy knowledge graphs. *CoRR*, abs/1812.00279, 2018.

[63] Guanglin Niu and Bo Li. Logic and commonsense-guided temporal knowledge graph completion, 2022.

[64] Guanglin Niu, Bo Li, Yongfei Zhang, and Shiliang Pu. Cake: A scalable commonsense-aware framework for multi-view knowledge graph completion, 2022.

[65] Guanglin Niu, Bo Li, Yongfei Zhang, Yongpan Sheng, Chuan Shi, Jingyang Li, and Shiliang Pu. Joint semantics and data-driven path representation for knowledge graph reasoning. *Neurocomputing*, 483:249–261, 2022.

[66] Guanglin Niu, Yongfei Zhang, Bo Li, Peng Cui, Si Liu, Jingyang Li, and Xiaowei Zhang. Rule-guided compositional representation learning on knowledge graphs. *CoRR*, abs/1911.08935, 2019.

[67] Simon Ott, Christian Meilicke, and Matthias Samwald. SAFRAN: an interpretable, rule-based link prediction method outperforming embedding models. *CoRR*, abs/2109.08002, 2021.

[68] Hiroki Ouchi, Jun Suzuki, Sosuke Kobayashi, Sho Yokoi, Tatsuki Kuribayashi, Ryuto Konno, and Kentaro Inui. Instance-based learning of span representations: A case study through named entity recognition. *CoRR*, abs/2004.14514, 2020.

[69] Matteo Paganelli, Paolo Sottovia, Francesco Guerra, and Yannis Velegrakis. Tuner: Fine tuning of rule-based entity matchers. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, CIKM '19, page 2945–2948, New York, NY, USA, 2019. Association for Computing Machinery.

[70] Matthew J Page, David Moher, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, Roger Chou, Julie Glanville, Jeremy M Grimshaw, Asbjørn Hróbjartsson, Manoj M Lalu, Tianjing Li, Elizabeth W Loder, Evan Mayo-Wilson, Steve McDonald, Luke A McGuinness, Lesley A Stewart, James Thomas, Andrea C Tricco, Vivian A Welch, Penny Whiting, and Joanne E McKenzie. Prisma 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372, 2021.

[71] Trishan Panch, Heather Mattie, and Rifat Atun. Artificial intelligence and algorithmic bias: implications for health systems. *Journal of global health*, 9(2), 2019.

[72] Pouya Pezeshkpour, Yifan Tian, and Sameer Singh. Investigating robustness and interpretability of link prediction via adversarial modifications. *CoRR*, abs/1905.00563, 2019.

[73] Forough Poursabzi-Sangdeh, Daniel G. Goldstein, Jake M. Hofman, Jennifer Wortman Vaughan, and Hanna M. Wallach. Manipulating and measuring model interpretability. *CoRR*, abs/1802.07810, 2018.

[74] Rehab Qarout, Alessandro Checco, Gianluca Demartini, and Kalina Bontcheva. Platform-related factors in repeatability and reproducibility of crowdsourcing tasks. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 7, pages 135–143, 2019.

[75] Kun Qian, Lucian Popa, and Prithviraj Sen. Systemer: A human-in-the-loop system for explainable entity resolution. 12(12):1794–1797, aug 2019.

[76] Meng Qu, Junkun Chen, Louis-Pascal A. C. Xhonneux, Yoshua Bengio, and Jian Tang. Rnnlogic: Learning logic rules for reasoning on knowledge graphs. *CoRR*, abs/2010.04029, 2020.

[77] Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *CoRR*, abs/1602.04938, 2016.

[78] Marco Tulio Ribeiro, Tongshuang Wu, Carlos Guestrin, and Sameer Singh. Beyond accuracy: Behavioral testing of nlp models with checklist. *arXiv preprint arXiv:2005.04118*, 2020.

[79] Tim Rocktäschel and Sebastian Riedel. End-to-end differentiable proving. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[80] Andrea Rossi, Donatella Firmani, Antonio Matinata, Paolo Merialdo, and Denilson Barbosa. Knowledge graph embedding for link prediction: A comparative analysis. *CoRR*, abs/2002.00819, 2020.

[81] Andrea Rossi, Donatella Firmani, Paolo Merialdo, and Tommaso Teofili. Explaining link prediction systems based on knowledge graph embeddings. In *Proceedings of the 2022 International Conference on Management of Data*, SIGMOD '22, page 2062–2075, New York, NY, USA, 2022. Association for Computing Machinery.

[82] Dongyu Ru, Changzhi Sun, Jiangtao Feng, Lin Qiu, Hao Zhou, Weinan Zhang, Yong Yu, and Lei Li. Learning logic rules for document-level relation extraction. *CoRR*, abs/2111.05407, 2021.

[83] Ali Sadeghian, Mohammadreza Armandpour, Patrick Ding, and Daisy Zhe Wang. DRUM: end-to-end differentiable rule mining on knowledge graphs. *CoRR*, abs/1911.00055, 2019.

[84] Cristina Sarasua, Elena Simperl, Natasha F Noy, Abraham Bernstein, and Jan Marco Leimeister. Crowdsourcing and the semantic web: A research manifesto. *Human Computation*, 2(1), 2015.

[85] Md Kamruzzaman Sarker, Lu Zhou, Aaron Eberhart, and Pascal Hitzler. Neuro-symbolic artificial intelligence. *AI Communications*, 34(3):197–209, 2021.

[86] Phillip Schneider, Tim Schopf, Juraj Vladika, Mikhail Galkin, Elena Simperl, and Florian Matthes. A decade of knowledge graphs in natural language processing: A survey. In *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 601–614, Online only, November 2022. Association for Computational Linguistics.

[87] August Th Schreiber, Guus Schreiber, Hans Akkermans, Anjo Anjewierden, Nigel Shadbolt, Robert de Hoog, Walter Van de Velde, and Bob Wielinga. *Knowledge engineering and management: the CommonKADS methodology*. MIT press, 2000.

[88] Gesina Schwalbe and Bettina Finzel. XAI method properties: A (meta-)study. *CoRR*, abs/2105.07190, 2021.

[89] Juan Sequeda and Ora Lassila. Designing and building enterprise knowledge graphs. *Synthesis Lectures on Data, Semantics, and Knowledge*, 11(1):1–165, 2021.

[90] Özge Sevgili, Artem Shelmanov, Mikhail Y. Arkhipov, Alexander Panchenko, and Chris Biemann. Neural entity linking: A survey of models based on deep learning. *CoRR*, abs/2006.00575, 2020.

[91] Hamed Shahbazi, Xiaoli Fern, Reza Ghaeini, and Prasad Tadepalli. Relation extraction with explanation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6488–6494, Online, July 2020. Association for Computational Linguistics.

[92] Vivian S. Silva, André Freitas, and Siegfried Handschuh. Exploring knowledge graphs in an interpretable composite approach for text entailment. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):7023–7030, Jul. 2019.

[93] Elena Simperl, Roberta Cuel, and Martin Stein. Incentive-centric semantic web application engineering. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 3(1):1–117, 2013.

[94] Elena Simperl and Markus Luczak-Rösch. Collaborative ontology engineering: a survey. *The Knowledge Engineering Review*, 29(1):101–131, 2014.

[95] Umutcan Simsek, Elias Kärle, Kevin Angele, Elwin Huaman, Juliette Opdenplatz, Dennis Sommer, Jürgen Umbrich, and Dieter Fensel. A Knowledge Graph Perspective on Knowledge Engineering. *SN Computer Science*, 4(1):16, October 2022.

[96] Rohit Singh, Venkata Vamsikrishna Meduri, Ahmed Elmagarmid, Samuel Madden, Paolo Papotti, Jorge-Arnulfo Quiané-Ruiz, Armando Solar-Lezama, and Nan Tang. Synthesizing entity matching rules by examples. *Proc. VLDB Endow.*, 11(2):189–202, oct 2017.

[97] Alison Smith-Renner, Ron Fan, Melissa Birchfield, Tongshuang Wu, Jordan Boyd-Graber, Daniel S. Weld, and Leah Findlater. No explainability without accountability: An empirical study of explanations and feedback in interactive ml. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–13, New York, NY, USA, 2020. Association for Computing Machinery.

[98] Robyn Speer, Joshua Chin, and Catherine Havasi. Conceptnet 5.5: An open multilingual graph of general knowledge. *CoRR*, abs/1612.03975, 2016.

[99] Josua Stadelmaier and Sebastian Padó. Modeling paths for explainable knowledge base completion. In *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 147–157, Florence, Italy, August 2019. Association for Computational Linguistics.

[100] Rudi Studer, V Richard Benjamins, and Dieter Fensel. Knowledge engineering: Principles and methods. *Data & knowledge engineering*, 25(1-2):161–197, 1998.

[101] Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, and Mariano Fernández-López. The neon methodology for ontology engineering. In *Ontology engineering in a networked world*, pages 9–34. Springer, 2011.

[102] Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. Yago: A core of semantic knowledge. In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, page 697–706, New York, NY, USA, 2007. Association for Computing Machinery.

[103] Haohai Sun, Jialun Zhong, Yunpu Ma, Zhen Han, and Kun He. Timetraveler: Reinforcement learning for temporal knowledge graph forecasting. *CoRR*, abs/2109.04101, 2021.

[104] Gytė Tamašauskaitė and Paul Groth. Defining a knowledge graph development process through a systematic review. *ACM Trans. Softw. Eng. Methodol.*, apr 2022. Just Accepted.

[105] Tommaso Teofili, Donatella Firmani, Nick Koudas, Vincenzo Martello, Paolo Merialdo, and Divesh Srivastava. Effective explanations for entity resolution models, 2022.

[106] Ilaria Tiddi and Stefan Schlobach. Knowledge graphs as tools for explainable machine learning: A survey. *Artificial Intelligence*, 302:103627, 2022.

[107] Trung-Kien Tran, Mohamed H. Gad-Elrab, Daria Stepanova, Evgeny Kharlamov, and Jannik Strötgen. Fast computation of explanations for inconsistency in large-scale knowledge graphs. In *Proceedings of The Web Conference 2020*, WWW '20, page 2613–2619, New York, NY, USA, 2020. Association for Computing Machinery.

[108] Samhita Vadrevu, Rakesh Nagi, JinJun Xiong, and Wen-mei Hwu. xER: An explainable model for entity resolution using an efficient solution for the clique partitioning problem. In *Proceedings of the First Workshop on Trustworthy Natural Language Processing*, pages 34–44, Online, June 2021. Association for Computational Linguistics.

[109] Michael van Bekkum, Maaike de Boer, Frank van Harmelen, André Meyer-Vitali, and Annette ten Teije. Modular design patterns for hybrid learning and reasoning systems: a taxonomy, patterns and use cases. *Applied Intelligence*, 51(9):6528–6546, 2021.

[110] Denny Vrandečić and Markus Krötzsch. Wikidata: A free collaborative knowledgebase. *Commun. ACM*, 57(10):78–85, sep 2014.

[111] Hailin Wang, Ke Qin, Guoming Lu, Jin Yin, Rufai Yusuf Zakari, and Jim Wilson Owusu. Document-level relation extraction using evidence reasoning on rst-graph. *Knowledge-Based Systems*, 228:107274, 2021.

[112] Hongwei Wang, Hongyu Ren, and Jure Leskovec. Relational message passing for knowledge graph completion. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, KDD '21, page 1697–1707, New York, NY, USA, 2021. Association for Computing Machinery.

[113] Jin Wang and Yuliang Li. Minun: Evaluating counterfactual explanations for entity matching. In *Proceedings of the Sixth Workshop on Data Management for End-To-End Machine Learning*, DEEM '22, New York, NY, USA, 2022. Association for Computing Machinery.

[114] Ping Wang, Khushbu Agarwal, Colby Ham, Sutanay Choudhury, and Chandan K. Reddy. Self-supervised learning of contextual embeddings for link prediction in heterogeneous networks. WWW '21, page 2946–2957, New York, NY, USA, 2021. Association for Computing Machinery.

[115] Xiaolong Wang, Yufei Ye, and Abhinav Gupta. Zero-shot recognition via semantic embeddings and knowledge graphs. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6857–6866, 2018.

[116] Xinru Wang and Ming Yin. Effects of explanations in ai-assisted decision making: Principles and comparisons. *ACM Trans. Interact. Intell. Syst.*, 12(4), nov 2022.

[117] Yuzhuo Wang, Hongzhi Wang, Junwei He, Wenbo Lu, and Shuolin Gao. Tagat: Type-aware graph attention networks for reasoning over knowledge graphs. *Knowledge-Based Systems*, 233:107500, 2021.

[118] Yuzhuo Wang, Hongzhi Wang, Wenbo Lu, and Yu Yan. Metranse: Manifold-like mechanism enhanced embedding for reasoning over knowledge graphs. *Expert Systems with Applications*, 209:118288, 2022.

[119] Gerhard Weikum, Luna Dong, Simon Razniewski, and Fabian M. Suchanek. Machine knowledge: Creation and curation of comprehensive knowledge bases. *CoRR*, abs/2009.11564, 2020.

[120] Matti Wiegmann, Michael Völske, Benno Stein, and Martin Potthast. Language models as context-sensitive word search engines. In *Proceedings of the First Workshop on Intelligent and Interactive Writing Assistants (In2Writing 2022)*, pages 39–45, Dublin, Ireland, May 2022. Association for Computational Linguistics.

[121] Christine T Wolf. From Knowledge Graphs to Knowledge Practices: On the Need for Transparency and Explainability in Enterprise Knowledge Graph Applications. 2020.

[122] Jianfeng Wu, Sijie Mai, and Haifeng Hu. Contextual relation embedding and interpretable triplet capsule for inductive relation prediction. *Neurocomputing*, 505:80–91, 2022.

[123] Junkang Wu, Wentao Shi, Xuezhi Cao, Jiawei Chen, Wenqiang Lei, Fuzheng Zhang, Wei Wu, and Xiangnan He. Disenkgat: Knowledge graph embedding with disentangled graph attention network. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, CIKM '21, page 2140–2149, New York, NY, USA, 2021. Association for Computing Machinery.

[124] Yi Xia, Mingjing Lan, Junyong Luo, Xiaohui Chen, and Gang Zhou. Iterative rule-guided reasoning over sparse knowledge graphs with deep reinforcement learning. *Information Processing & Manage-*

ment, 59(5):103040, 2022.

[125] Yuxin Xiao, Zecheng Zhang, Yuning Mao, Carl Yang, and Jiawei Han. SAIS: supervising and augmenting intermediate steps for document-level relation extraction. *CoRR*, abs/2109.12093, 2021.

[126] Qizhe Xie, Xuezhe Ma, Zihang Dai, and Eduard H. Hovy. An interpretable knowledge transfer model for knowledge base completion. *CoRR*, abs/1704.05908, 2017.

[127] Wenhan Xiong, Thien Hoang, and William Yang Wang. Deeppath: A reinforcement learning method for knowledge graph reasoning. *CoRR*, abs/1707.06690, 2017.

[128] Vikas Yadav and Steven Bethard. A survey on recent advances in named entity recognition from deep learning models. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2145–2158, Santa Fe, New Mexico, USA, August 2018. Association for Computational Linguistics.

[129] Fan Yang, Zhilin Yang, and William W. Cohen. Differentiable learning of logical rules for knowledge base completion. *CoRR*, abs/1702.08367, 2017.

[130] Zijun Yao, Chengjiang Li, Tiansi Dong, Xin Lv, Jifan Yu, Lei Hou, Juanzi Li, Yichi Zhang, and Zelin Dai. Interpretable and low-resource entity matching via decoupling feature learning from decision making. *CoRR*, abs/2106.04174, 2021.

[131] Hongbin Ye, Ningyu Zhang, Hui Chen, and Huajun Chen. Generative knowledge graph construction: A review. *CoRR*, abs/2210.12714, 2022.

[132] Jinyoung Yeo, Haeju Park, Sanghoon Lee, Eric Wonhee Lee, and Seung-won Hwang. Xina: Explainable instance alignment using dominance relationship. *IEEE Transactions on Knowledge and Data Engineering*, 32(2):388–401, 2020.

[133] Xu Yuan, Qihang Lei, Shuo Yu, Chengchuan Xu, and Zhikui Chen. Fine-grained relational learning for few-shot knowledge graph completion. *SIGAPP Appl. Comput. Rev.*, 22(3):25–38, nov 2022.

[134] Shuang Zeng, Yuting Wu, and Baobao Chang. SIRE: separate intra- and inter-sentential reasoning for document-level relation extraction. *CoRR*, abs/2106.01709, 2021.

[135] Canlin Zhang, Chun-Nan Hsu, Yannis Katsis, Ho-Cheol Kim, and Yoshiki Vázquez-Baeza. Theoretical rule-based knowledge graph reasoning by connectivity dependency discovery. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–9, 2022.

[136] Wen Zhang, Shumin Deng, Mingyang Chen, Liang Wang, Qiang Chen, Feiyu Xiong, Xiangwen Liu, and Huajun Chen. Knowledge graph embedding in e-commerce applications: Attentive reasoning, explanations, and transferable rules. IJCKG'21, page 71–79, New York, NY, USA, 2022. Association for Computing Machinery.

[137] Wen Zhang, Shumin Deng, Han Wang, Qiang Chen, Wei Zhang, and Huajun Chen. Xtranse: Explainable knowledge graph embedding for link prediction with lifestyles in e-commerce. In Xin Wang, Francesca A. Lisi, Guohui Xiao, and Elena Botoeva, editors, *Semantic Technology*, pages 78–87, Singapore, 2020. Springer Singapore.

[138] Wen Zhang, Bibek Paudel, Wei Zhang, Abraham Bernstein, and Huajun Chen. Interaction embeddings for prediction and explanation in knowledge graphs. WSDM '19, page 96–104, New York, NY, USA, 2019. Association for Computing Machinery.

[139] Yongqi Zhang and Quanming Yao. Knowledge graph reasoning with relational directed graph. *CoRR*, abs/2108.06040, 2021.

[140] Yuhui Zhang, HAO DING, Zeren Shui, Yifei Ma, James Zou, Anoop Deoras, and Hao Wang. Language models as recommender systems: Evaluations and limitations. In *I (Still) Can't Believe It's Not Better! NeurIPS 2021 Workshop*, 2021.

[141] Wenxuan Zhou, Hongtao Lin, Bill Yuchen Lin, Ziqi Wang, Junyi Du, Leonardo Neves, and Xiang Ren. Nero: A neural rule grounding framework for label-efficient relation extraction. In *Proceedings of The Web Conference 2020*, WWW '20, page 2166–2176, New York, NY, USA, 2020. Association for Computing Machinery.

[142] Anjie Zhu, Deqiang Ouyang, Shuang Liang, and Jie Shao. Step by step: A hierarchical framework for multi-hop knowledge graph reasoning with reinforcement learning. *Knowledge-Based Systems*, 248:108843, 2022.

[143] Unai Zulaika, Aitor Almeida, and Diego López-de Ipiña. Influence functions for interpretable link prediction in knowledge graphs for intelligent environments. In *2022 7th International Conference on Smart and Sustainable Technologies (SpliTech)*, pages 1–7, 2022.

[144] Andrew Zupon, Maria Alexeeva, Marco Valenzuela-Escárcega, Ajay Nagesh, and Mihai Surdeanu. Lightly-supervised representation learning with global interpretability. In *Proceedings of the Third*

Workshop on Structured Prediction for NLP, pages 18–28, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.