Ethical and Sustainability Considerations for Knowledge Graph based Machine Learning

Carsten Felix Draschner
Information Systems and Artificial Intelligence
University of Bonn
Bonn, Germany
0000-0002-1006-146X

Hajira Jabeen
Big Data Analytics
Leibniz GESIS Institute
Cologne, Germany
0000-0003-1476-2121

Jens Lehmann
Amazon
(Work done while at University of Bonn.)
Dresden, Germany
0000-0001-9108-4278

***Abstract*—Artificial Intelligence (AI) and Machine Learning (ML) are becoming common in our daily lives. The AI-driven processes significantly affect us as individuals and as a society, spanning across ethical dimensions like discrimination, misinformation, and fraud. Several of these AI & ML approaches rely on Knowledge Graph (KG) data. Due to the large volume and complexity of today's KG-driven approaches, enormous resources are spent to utilize the complex AI approaches. Efficient usage of the resources like hardware and power consumption is essential for sustainable KG-based ML technologies. This paper introduces the ethical and sustainability considerations, challenges, and optimizations in the context of KG-based ML. We have grouped the ethical and sustainability aspects according to the typical Research & Development (R&D) lifecycle: an initial investigation of the AI approach's responsibility dimensions; technical system setup; central KG data analytics and curating; model selection, training, and evaluation; and final technology deployment. We also describe significant trade-offs and alternative options for dedicated scenarios enriched through existing and reported ethical and sustainability issues in AI-driven approaches and research. These include, e.g., efficient hardware usage guidelines; or the trade-off between transparency and accessibility compared to the risk of manipulability and privacy-related data disclosure. In addition, we propose how biased data and barely explainable AI can result in discriminating ML predictions. This work supports researchers and developers in reflecting, evaluating, and optimizing dedicated KG-based ML approaches in the dimensions of ethics and sustainability.**

***Index Terms*—Semantic Processing, Knowledge Graphs, Machine Learning, AI Ethics, Sustainable Machine Learning, Explainable AI, RDF**

I. Introduction

Artificial Intelligence (AI) implemented through Machine Learning (ML) driven processes increasingly impacts our daily lives. Those reach from how we: buy products, consume streaming content, preselect CVs in job applications, and interact with customer service over chatbots. A significant share of these scenarios relies on Knowledge Graph (KG) data [1]. The Semantic Web [2] vision describes the joint effort to integrate various internet data sources into a relation-centric linked data structure. Out of this concept, several open KGs have emerged (DBpedia [3], YAGO [4], Freebase [5], Wikidata [6]) as well as big tech companies implemented enterprise KGs for their products (Meta, Google, LinkedIn, eBay, IBM) [1]. The KG-based ML models have already substantially impacted individuals and overall society. Based on already reported problems and challenges with AI-driven approaches [7]–[12], we see a need to investigate KG-based ML. Due to the complexity of the applied ML models and the enormous underlying training data, the investigation and optimization of respective processes are challenging. So we need best practice guidelines to offer Data Scientists and ML Engineers a starting point for improved and more holistic KG ML R&D. Within this work, we provide a set of novel contributions:

- Introduction of a novel Ethical AI perspective on KG-based ML.
- Sustainability considerations on the KG-based ML R&D lifecycle.
- Presentation of interwoven ethical and sustainability dimensions along the Downstream Pipeline, including Responsible AI Approach, Technical Setup, Data Insights, ML, Training & Evaluation, and finally, Deployment.
- Ethical and Sustainable KG-based ML pipeline components, each structured by Definition, Challenges, Examples, Research Questions, and Recommendations.

Section II introduces the major terms: Ethical AI, Sustainable AI, Knowledge Graphs, AI/ML, and KG-based ML. In Section III, we introduce the classical schema of a KG-based ML R&D Pipeline. In Sections IV-IX, we introduce for each step of this Pipeline Ethical and Sustainability considerations for KG-based ML. Within the Conclusion, in Section X, we summarize our work and provide future work directions.

II. Preliminaries

A. Knowledge Graphs

Knowledge Graphs (KGs) are linked data representations of information. Entities and values are associated with directed relations. A KG corresponds to a directed labeled multi-graph.

B. Artificial Intelligence, and Machine Learning

Artificial intelligence is the theory and development of computer systems able to perform tasks usually requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages [13]. One option to reach and implement AI is Machine Learning (ML). ML is the capacity of computers to learn and adapt without following explicit instructions by using algorithms and statistical models to analyze and infer from patterns in data [14]. Despite the conceptual differences, AI and ML are often used synonymously.

C. Knowledge Graph based Machine Learning

In several approaches, KGs are the key element of ML pipelines. On the one hand, ML approaches can perform knowledge retrieval to enrich a KG database. In those cases, information is extracted from multi-modal data sources like natural language texts, images, tables, audio, and video. Also, existing KG can be further enriched over Link Prediction, Entity Resolution, Entity Matching, and Inference. On the other hand, KG can be used as a source for downstream ML pipelines. Those pipelines can implement Classification, Regression, or Recommendation Systems. The ML model uses stored values inside the KG and the graph structure. As the structure can be arbitrarily complex, and common ML models need fixed-length numeric feature tensors across all samples, KG-based ML uses latent embeddings to reach through, e.g., Knowledge Graph Embeddings (KGE). Especially the second case is considered to be KG-based ML, while the initial creation and enrichment of the KG can also be a significant source of ethical implications.

D. Artificial Intelligence Ethics

AI Ethics describes the efforts to investigate ethical and moral aspects of AI-driven processes [11], [15]. As AI-driven processes can impact our daily lives and might perform autonomous decisions based on pre-trained models, those can lead to unethical, discriminating situations [11]. Bad AI behavior can be due to various reasons, like data bias or model failure. Especially complex neural networks are perceived as black-box systems. Through the effort of Explainable AI, the models are leveraged to create reasonable and explainable results [8], [11]. A typical example of Ethical AI comes from the already known trolley dilemma. This example is related to the decision-making process of self-driving cars in an unavoidable accident situation where an autonomous system needs to decide between several evils [15]–[17].

E. Sustainable Artificial Intelligence

Sustainable AI includes optimizations of AI/ML methods that minimize the resource consumption of ML algorithms and, if necessary, also justify it concerning the purpose [7], [8], [18]. In contrast to Sustainable AI, AI for Sustainability are approaches that use AI to drive processes that lead to a more sustainable world [7].

III. KG based ML Pipeline Schema

The R&D lifecycle of KG-based ML follows the classical scheme in Data Science project pipelines. In the following sections, we present considerations for each step that investigate both the sustainability and ethical dimensions. The first step, Responsible AI Approach (see Sect. IV), introduces the initial use case from which the KG-based approach derives. In the next step, the needed technical System Setup (see Sect. V) is created for later deployment. Subsequently, the Data Insides are explored, curated, and processed to serve as suitable training and evaluation data (see Sect. VI). An ML model is selected and further developed that fits the use case and the available data VII. This model is optimized by iterative Training and Evaluation processes to achieve the targeted performance (see Sect. VIII). In the final Deployment, the models and the results are made available to the users for further development, reproducibility, and interaction (see Sect. IX). Each of these steps includes dedicated optimizations for Sustainable AI and AI Ethics in general but also requires a dedicated focus on the particular issues when dealing with KG-based ML. Fig. 1 presents the respective steps according to related research questions.

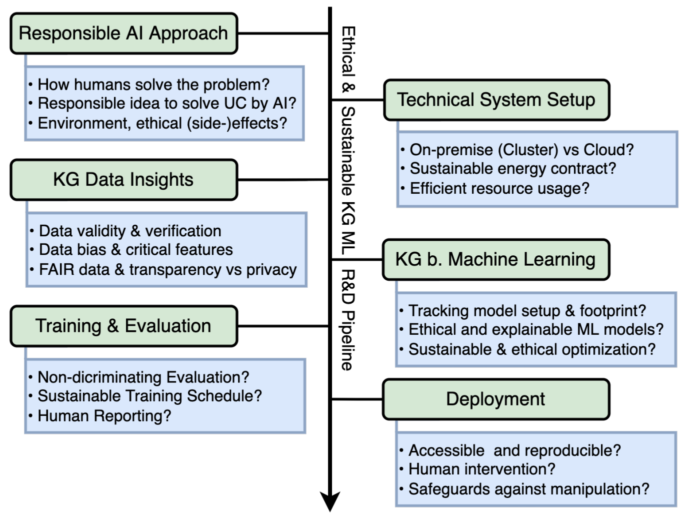

Fig. 1: Ethical and Sustainable KG based ML pipeline

IV. Responsible AI Approach

Even before any technical decision has been made, any line of code has been written, the multidisciplinary team of stakeholders should reflect on the ethical and sustainable implications of the technical KG-driven ML approach and should have answers to the following questions: What are the foreseeable edge cases, what are side effects? To what extent will a tool be developed that allows unintended dual use? How has the problem been solved in the past by humans in a non-AI-driven process.?

A. Ethical KG based ML Approach

- ***Definition:*** In the first step, one must reflect on the ethical implications and side effects of a KG-based ML approach [8], [11].
- ***Challenges:*** It is impossible to conclude all different effects directly from the initial application. In basic research, complex models are trained and evaluated using illustrative examples, which can then be applied to more

socially relevant topics. Same time, the tremendous integration opportunities of data integration and the chances of conversational AI can lead to unintended dual use of technology [19].

- ***Examples:*** Chatbots handle customer service centers for e-commerce products. However, chatbots do not take emergency calls, although both implement dialog-based KG-based conversational AI systems to ask for help and solve problems [15].
- ***Research Questions:*** Have the social and ethical implications of the specific KG-driven AI been evaluated and allowed for responsible research and development of this technology?
- ***Recommendations:*** The ethical implications of a KG-driven AI process should be assessed and justified as an initial step. The stakeholders concerned can initiate a multi-layered evaluation in multidisciplinary teams. It can reflect how it was solved in the past by humans or classical imperative programs and if the use case should be solved autonomously at all [15].

B. *Responsible Investment of Resources*

- ***Definition:*** Applying AI, in particular, to KG can be processing intensive which has sustainability implications. [7], [8], [11], [20]
- ***Challenges:*** It is difficult to justify the significant power and hardware resources directly through use cases. Trained KGE corpora can indirectly add value as fundamental research, and created resources offer amortization in follow-up approaches.
- ***Examples:*** The AI Alpha Go Zero, which is optimized to play a well-known board game GO, generated 96 tons of CO2 in 40 days of research training, equivalent to 1000h of air travel or 23 American households [7].
- ***Research Questions:*** Are the resources used worth consuming? Are transfer learning use cases foreseeable that validate the spending resources being reasonable? Is it possible to develop to showcase KG-based AI development directly on Use Cases with social or environmental positive impact? Is a KG the most practical and energy-efficient data representation for the dedicated follow-up AI approaches?
- ***Recommendations:*** Develop and showcase KG-based AI development directly on Use Cases with social or environmental positive impact. Sample KG AIs should be accompanied by accessible and reusable KG data s.t. invested resources can more likely amortize over multiple projects.

V. System Setup

As KGs can be huge and processing, those have direct implications on needed hardware complexity but also foreseeable runtime. Both the hardware and electricity footprint should be minimized under sustainability considerations. In KG querying and embeddings, the whole data should be quickly accessible, implying the need for distributed architectures. On the other hand creation of KGE need synchronized instances of embeddings which is a technical bottleneck.

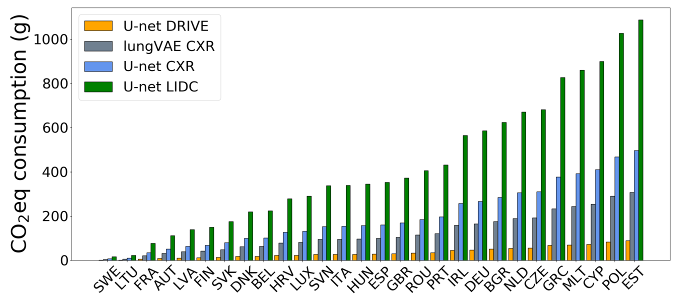

Fig. 2: Estimated CO2 emissions of Carbon Tracker experiment [9]

A. *Efficient Hardware Usage*

- ***Definition*:** KG-based AI R&D lifecycle is performed on computers. The necessary resources have a sustainability footprint [9], [20]. This consists of the acquisition and production of the hardware, the power consumption, and the expected longevity [7], [21].
- ***Challenges*:** The hardware capabilities of KG-based ML can be immense as KGE correlate to large language model training that is processing power and memory intensive [20]. The selection of hardware involves trade-offs between reusing existing hardware or accumulating multiple nodes in a cluster to make the necessary processing power available. New hardware may still be essential as requirements increase because it has more processing power or features and is more power-efficient [21]. In addition, specific applications require dedicated hardware and software platforms. With on-premise system architecture, more individual decisions can be made than with cloud computation systems [21], while possible underutilization barely justifies the initial hardware costs [22].
- ***Examples*:** (1) KGE needs to be trained once for one Dataset, and from then on, it can be further used until KG data has been updated. With Cloud Computing, PaaS, and AI as a service, the hardware can be used more efficiently over the whole day and year. Fewer instances of bought hardware need to stay idle if load balancing and sharing economy on servers are used. (2) If existing hardware can be combined in clusters through distributed computing architectures, existing resources can be used longer while more processing power is needed
- ***Research Questions*:** How can the hardware resources be used, reused, and utilized as efficiently as possible?
- ***Recommendations*:** Hardware should be deployed and efficiently utilized either as a cloud solution or alternatively as a crowd-shared on-premise architecture. Software abstraction layers and containerization should contribute to hardware-independent R&D and comfortable deployment.

B. Optimizing Carbon Footprint

- ***Definition*:** As the on-premise or cloud systems need electricity majorly as an ongoing resource, the footprint also correlates to the electric energy production footprint [7], [9], [21]. KG-based ML, e.g., using KGE, needs regular retraining as KGs are changing, and corresponding models must be optimized to capture changes in data structure and content.
- ***Challenges*:** One does not always have direct access to the power contract of cloud or on-premise systems to optimize it in the sustainability dimension. Also, the hardware produces heat. As KG-based AI can be processing intensive, the energy used should be minimized and fed by renewable energy sources. [20]–[22]
- ***Examples*:** The European CO2 footprint of electricity production differs significantly due to a different share of renewable energies, nuclear power, and fossil fuels [23]. Because of weather and solar radiation, the CO2 footprint is also dependent on the time of day (Figs. 2 & 6).
- ***Research Questions*:** How can the resource carbon footprint be minimized? Is it possible to schedule the regular KG-based ML processes in times when more sustainable energy is available (see Fig. 2)?
- ***Recommendations*:** The hardware used should be based on CO2-neutral energy sources [21]. As part of the setup, tracking [9] and documentation of the effectively used resources and emissions footprint create transparency. Schedule regular resource-intensive KG-based ML processes like the KGE optimization being executed when more renewable energy is available (see Fig. 6).

VI. Data Insights

The KG data specifics play an extraordinary role in the performance and behavior of KG-based AI processes [24]. The initial data selection decides which sources of information and data are accessible to the process. Existing KGs can be selected, merged, or further enriched by non-KG data.

Fig. 3: Google Search KG result about actor and pop-figure "Chuck Norris", see what is stated as **Fact**, Example taken from Vang et. al. [10]

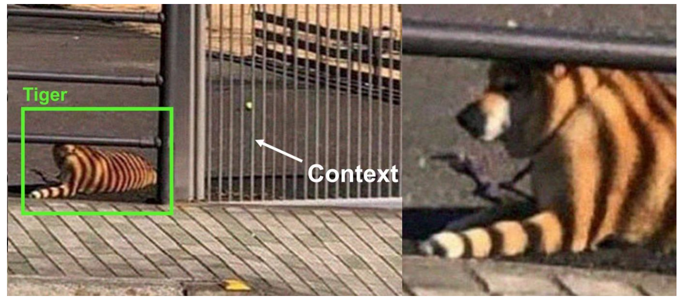

Fig. 4: Sample Image [25] motivating context in ML data

A. Knowledge Graph Data Reliability

- ***Definition*:** KGs are created and enriched by a variety of processes. Incorrect or inaccurate information can be mapped through the source data and process. Also, the concept of open-world assumption (OWA) is part of KG data. OWA states that information can be correct even if this data is not in KG.
- ***Challenges*:** The sources of false information are complex and sometimes difficult to trace. False data is not always apparent at first sight.
- ***Examples*:** (1) Automatic retrieval of KG data from texts or other multi-modal data can create inaccurate knowledge data as those techniques do not work perfectly [26]. (2) The Open World Assumption (OWA) is a central element in KG data, implying that information can also exist if not stored in the KG. This OWA idea implies causality if a KG does not contain specific data; this information can still be accurate, which can lead by chance to problems in the negative sampling in KGEs. (3) Also, fraud and manipulation can happen as, e.g., DBpedia extracts its information from Wikipedia, but public figures and institutions try to influence how they are perceived and presented on the internet [8]. (4) Not only bad intention but also humor or sarcasm can lead to incorrect data as facts about pop culture people, e.g., Chuck Norris (see Fig. 3) [10]. (5) As facts change, KG data could also be not up to date [27].
- ***Research Questions*:** What is the data quality, especially to which extent is the stored information reliable? With which intent, by whom, and through which processes has this data been created?
- ***Recommendations*:** Quality, noise, and incorrect data are identified through data analytics and should be documented and published through transparent FAIR principles.

B. Protecting Privacy vs. Enrichment of Context

- ***Definition*:** KGs have great potential to combine many data sources and make them accessible to humans and machines. More context can support AI-driven predictions (see Fig. 4). For trustworthy AI, underlying transparent and FAIR training data and metadata are essential.

- ***Challenges*:** However, a high level of transparency of fully integrated data can also be exploited and used for ethical purposes [19]. Therefore, there is a trade-off in data accessibility between high transparency and the need-to-know principle [28]. The model can be more challenging to reproduce and validate with reduced transparency by independent institutions.
- ***Examples*:** In some countries, personal relationships or research in specific subject areas can lead to discrimination, persecution, or prison [19].
- ***Research Questions*:** For what ethical purposes can transparent data be used? How can data be made both sufficiently transparent for validation and securely accessible? Which training and metadata should be made accessible to humans and machines?
- ***Recommendations*:** Take a clear position in the trade-off between the need-to-know and transparency. Enable independent validation of technologies to justify why the transparently available data is likely safe [8], [11].

VII. Knowledge Graph-based Machine Learning

KG-based ML is a major source of various AI approaches, including Conversational AI, Question Answering Systems, Recommendation Systems, and many more. The data integration opportunities of KG as a central data source for training and result semantification offer high expressivity of multi-modal linked knowledge. With the sheer size and restrictions of current ML approaches, challenges are given in dimensions of explainability and reproducibility.

A. Ethical & Explainable ML for KGs

- ***Definition*:** KG-based ML models and the corresponding latent embeddings are multidimensional parameters and feature spaces encoding properties and processes of the AI. Accessibility, reproducibility, reusability, and explainability are essential for ethical AI [11], [15], [29].
- ***Challenges*:** Since the data and features in KG are arbitrarily complex for each sample, and standard ML models assume fixed numeric feature vectors, latent embeddings like KGE are necessary [20], [30], [31]. Especially the high-performance neural network models and latent embeddings significantly complicate explainability and lead to AI often being perceived as a black box [32]–[34].
- ***Examples*:** (1) The approach distilling neural networks transforms complex neural network ML models into more explainable decision forest models [35]. (2) Explainable AI approaches can also describe the features or embedding components that were most influential [8] for a specific prediction.
- ***Research Questions*:** Can the models produce explainable predictions? Does KG-based ML's complexity also allow less powerful hardware to use the new technology?
- ***Recommendations*:** Use existing benchmarks to fit already developed models instead of developing new ones [23]. Use KG-based ML models that focus on explainability. With conversational AI and result semantification, the results become intuitively more accessible [8], [36]. Use scalable models for KG-based AI [37]–[39] also working on distributed systems [30], [31].

B. Critical and Sensitive Features

- ***Definition*:** KG ML training is based on structural and value features. The distribution of the values may not represent the accurate distribution of the values [15]. Also, special features are associated with discrimination, so handling these features poses particular challenges [19], [40].
- ***Challenges*:** The simple omission/deletion of features can still lead to discrimination because, in these cases, meta-data may still allow unintended feature reconstruction [12].
- ***Examples*:** Features associated with discrimination include age, gender, heritage, skin color, political orientation, sexual orientation, and religion [8].
- ***Research Questions*:** Which features associated with discrimination are present in the KG? In which dimensions is there a bias, and why? Can features be aggregated, pseudonymized, or re-balanced to reduce bias?
- ***Recommendations*:** Identify the presence of critical features. Describe bias distributions in these features. Aggregate, pseudonymize or remove critical features in a transparent process. Evaluate whether removed features can be reconstructed through metadata. Optimize labeling and curation procedures to minimize bias [27], [41].

VIII. Training and Evaluation

Significant problems of unfair AI predictions result from suboptimal training. Skewed training data, unintended optimization strategies resulting in discriminating optima lead to results which can be later challenging to explain or fix due to the complexity of nowadays embeddings based and neural network ML pipelines. Same time the training of ML models is a fundamental part of spent resources within the KG-based ML R&D lifecycle. So training and evaluation should be optimized to reduce the carbon footprint.

Fig. 5: BBC article [42] about hurtful/racism ML annotated Google Photos image category

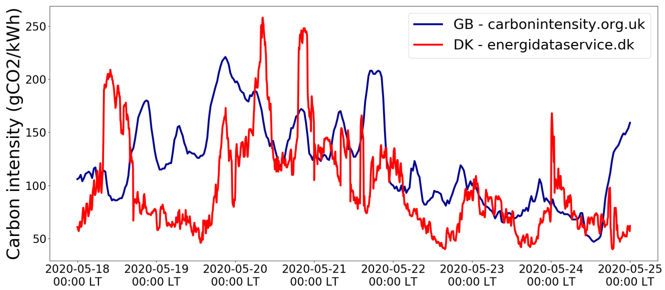

Fig. 6: Real-time carbon intensity (gCO 2eq/kWh) for Denmark (DK) and Great Britain (GB) from 2020-05-18 to 2020-05-25 shown in local time from Carbon Tracker analytics [9]

A. Bias in Training and Evaluation

- ***Definition*:** ML models are trained by minimizing the error of predicted results compared to the actual annotated true label. The performance of a model is measured by performance over unseen samples.
- ***Challenges*:** If the test and validation data are biased or do not contain sufficient data about possibly discriminated entities, they will not be present in overall precision or recall measures [11].
- ***Examples*:** (1) ML-based image grouping within Google Photos led to a hurtful, racist classification of people of colour [42]. (2) The Microsoft chatbot Tay was influenced by Twitter users to state offending and hate speech texts (See Fig. 5 & Fig. 7) [43].
- ***Research Questions*:** Are the test and validation set biased? Are opportunities available to live annotated in production KG AI predictions?
- ***Recommendations*:** A definition of critical samples and tests of dedicated edge cases should be performed. Users or affected humans should have the option to report problematic predictions in live systems as it is already typical for bug and issue reports in open source projects [8], [27]. The model should be tested before deploying and afterward to ensure that the initially made claims still suit outside world scenarios [8]. As part of evaluation dissemination, the performance across all samples and subset performances should be accessible. Additionally, the performance across critical features should be validated [12]. With the opportunities for data integration of KGs, an improved opportunity is given to identify bias among samples and optimize the training data distribution.

B. Sustainable Training

- ***Definition*:** Training KG-based ML requires considerable resources since both latent embeddings and (Graph-) Neural Networks have significant training efforts [9], [30], [31]. During the initial exploration of the models, the optimal configuration of the hyper-parameters must be found.
- ***Challenges*:** Since KGE are optimized from random vectors and represent exact KG entities, these models cannot handle out-of-sample entities by default. Models that support out-of-sample handling or inductive link prediction require further adjustments to minimize retraining from scratch [20]. Additionally, the search within the hyper-parameter grid requires recurrent training of the same model.
- ***Examples*:** The training of latent embeddings KGE exceeds the already vast complexity of large language models [20], [30] which already produced huge carbon emissions [44], [45] (see Table I).
- ***Research Questions*:** How does KG-based AI deal with samples not yet seen? How can transfer learning and update ability be implemented in the approaches so that it is not necessary to train from scratch with changing KG data or novel entities [20]?
- ***Recommendations*:** Evaluate the KG data volatility. Use KG-based AI models that allow out-of-sample learning [46], inductive link prediction, and transfer learning. Reduce the hyper-parameter space to a minimum in a grid search and use early stopping to minimize overfitting and unnecessary training cycles [9], [21]. Further data and model compression optimization can additionally minimize the necessary resources [38].

IX. Accessible Deployment

In general, several optimizations targeting accessible deployment for AI are also applicable to KG-based ML. These include easy accessibility over good documentation, open source code, sample notebooks, AIaaS, and a peer-reviewed scientific publication. These improve the trust, reusability, and reproducibility of the approach. The opportunities of result semantification offer opportunities to enhance ethical and fair KG-based ML behavior.

Scenario	Carbon Footprint
Roundtrip flight b/w NY and SF (one passenger)	1,984
Human life (avg. one year)	11,023
American life (avg. one year)	36,156
US car including fuel (avg. one lifetime)	126,000
Transformer (213M parameters) W/ neural architecture search	626,155

TABLE I: Carbon Footprint Large Language Model [44], [45]

A. Prediction and Meta Data Semantification

- ***Definition*:** For the application, validation, and further development of KG-based AI, accessibility and reproducibility are fundamentally important. In addition to the technical details, resource consumption can also be tracked [7], [9].
- ***Challenges*:** Throughout the ML pipeline, there are many (hyper-)parameters and resource-consuming instances necessary for holistic tracking. At the same time, the carbon footprint is not traceable at all times since not every phase transparently breaks down the used resources.
- ***Examples*:** The tool Carbontracker supports the automated evaluation of CO2 footprints [7], [9], [21], [45].

Fig. 7: Blog post from the Verge article [43] about manipulated Chatbot AI stating offending and hate speech texts

- ***Research Questions***: Which resources were used throughout the KG ML lifecycle, and how can they be tracked automatically? Which (hyper-)parameters and configuration information are necessary to use, reproduce, validate and further develop the entire KG AI [7], [11]?
- ***Recommendations***: Existing AI CO2 tracking systems can be used to add resource consumption to documentation automatically (see Fig. I) [7], [9], [21], [45]. The use of standardized ontologies like MEX and MLschema allows a human and machine-readable tracking and documentation of the AI setup [47], [48].

B. Crowd Sourced AI Labeling and Intervention

- ***Definition***: AI trends like Conversational AI also rely on KG-based ML. These allow both machine and human-readable ML data and predictions.
- ***Challenges***: ML models are optimized on training, test, and validation sets, but after they are deployed, these models will also face unseen samples. The resulting predictions can be wrong or perceived to be unethical. The interaction with ML in live systems also allows manipulation and fraud. Same time, the exchange can also have an underlying bias in contributors [41].
- ***Examples***: The Never Ending Learning Project NELL enriches its knowledge base and offers the crowd to interact with recent KG-based ML results improving the ML performance and validity of knowledge base [27].
- ***Research Questions***: Does the KG-based ML provide reporting and intervention options to improve the KG base and its ML predictions? What are the safeguards against manipulation and bias of crowd-sourced contributors [8]?
- ***Recommendations***: Provide interaction and reporting mechanisms for crowd-sourced interaction with the AI [27]. Track and report the influence of crowd to ML predictions [8]. Generate accessible and understandable ML predictions aligned to conversational and explainable AI concepts. Benchmark the developed KG-based ML with critical unseen examples focused on non-discriminating and ethical model performance.

X. Conclusion

The opportunities for ethical and sustainability optimizations for KG-based ML are manifold. This work introduced the most prominent ethical concerns raised by the large-scale deployment of KG-based ML applications. We show how sustainability, ethics, KG, and ML are interwoven across the entire R&D life cycle. The possibilities range from an initial reflective check on how far use cases should be automated by AIs, taking into account potential side effects and edge cases. Even during the initial setup of the processing environment, foreseeable resource utilization can be optimized and technical reproducibility can be improved. The data in KG-based ML pipelines also significantly affect the process and should be evaluated regarding reliability, bias, and fairness. In addition, careful handling of special features is essential, including the trade-off between sufficient data context and privacy principles. The development of ML models can also be optimized technically in pipeline development through more sustainable training and ethical models with the help of explainable AI, fair evaluation, and accessible deployment. The deployed models should fulfill a broad range of interaction options with humans to report problems, handle unseen samples, and be used in transfer learning tasks. We introduce the application of AI Ethics and Sustainable AI to the KG-based ML domain. The work presented here is not intended to be complete as the field is broad. However, it is a starting point for KG-based AI R&D teams interested in optimizing technology under ethical and sustainability concerns. The findings are presented by a transfer of ethical, sustainable, and novel introduced considerations, including examples of past problems and suggestions for hands-on optimizations and solutions.

A. Future Work

The concepts presented here are mostly part of every KG-based ML pipeline. We believe that through easy-to-use tools and integration into widely used libraries, assistance should be offered as already the first applications provide it [9]. We are committed to supporting the process of ethical and sustainable KG-based ML research and development with hands-on solutions to advance this field. Corresponding to the FAIR concept, this paper should introduce meta dimensions of good research and development for KG-based ML in dimensions of Ethics and Sustainability.

Acknowledgements

This work was partly supported by the EU Horizon 2020 project PLATOON (Grant agreement ID: 872592).

References

[1] N. Noy, Y. Gao, A. Jain, A. Narayanan, A. Patterson, and J. Taylor, "Industry-scale knowledge graphs: lessons and challenges," *Communications of the ACM*, vol. 62, no. 8, pp. 36–43, 2019.

[2] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific american*, vol. 284, no. 5, pp. 34–43, 2001.

[3] J. Lehmann, R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P. N. Mendes, S. Hellmann, M. Morsey, P. Van Kleef, S. Auer *et al.*, "Dbpedia–a large-scale, multilingual knowledge base extracted from wikipedia," *Semantic web*, vol. 6, no. 2, pp. 167–195, 2015.

[4] F. M. Suchanek, G. Kasneci, and G. Weikum, "Yago: a core of semantic knowledge," in *Proceedings of the 16th international conference on World Wide Web*, 2007, pp. 697–706.

[5] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor, "Freebase: a collaboratively created graph database for structuring human knowledge," in *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, 2008, pp. 1247–1250.

[6] D. Vrandečić and M. Krötzsch, "Wikidata: a free collaborative knowledgebase," *Communications of the ACM*, vol. 57, no. 10, pp. 78–85, 2014.

[7] A. van Wynsberghe, "Sustainable ai: Ai for sustainability and the sustainability of ai," *AI and Ethics*, pp. 1–6, 2021.

[8] L. Floridi, J. Cowls, T. C. King, and M. Taddeo, "How to design ai for social good: Seven essential factors," in *Ethics, Governance, and Policies in Artificial Intelligence*. Springer, 2021, pp. 125–151.

[9] L. F. W. Anthony, B. Kanding, and R. Selvan, "Carbontracker: Tracking and predicting the carbon footprint of training deep learning models," *arXiv preprint arXiv:2007.03051*, 2020.

[10] K. J. Vang, "Ethics of google's knowledge graph: some considerations," *Journal of Information, Communication and Ethics in Society*, 2013.

[11] S. L. Piano, "Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward," *Humanities and Social Sciences Communications*, vol. 7, no. 1, pp. 1–7, 2020.

[12] J. Dustin. (2018) Amazon scraps secret ai recruiting tool that showed bias against women. [Online]. Available: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

[13] O. E. Dictionary. (2022) artificial intelligence. [Online]. Available: https://www.oed.com/viewdictionaryentry/Entry/271625

[14] ——. (2022) Machine learning. [Online]. Available: https://www.oed.com/view/Entry/111850?redirectedFrom=machine+learning+#eid38479194

[15] D. Greene, A. L. Hoffmann, and L. Stark, "Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning," in *Proceedings of the 52nd Hawaii international conference on system sciences*, 2019.

[16] N. J. Goodall, "Ethical decision making during automated vehicle crashes," *Transportation Research Record*, vol. 2424, no. 1, pp. 58–65, 2014.

[17] H. Ye, L. Liang, G. Y. Li, J. Kim, L. Lu, and M. Wu, "Machine learning for vehicular networks: Recent advances and application examples," *ieee vehicular technology magazine*, vol. 13, no. 2, pp. 94–101, 2018.

[18] M. Agravante. (2020) Mit moves toward greener, more sustainable artificial intelligence. [Online]. Available: https://inhabitat.com/mit-moves-toward-greener-more-sustainable-artificial-intelligence/

[19] R. Huber and J. Klump, "The dark side of the knowledge graph-how can we make knowledge graphs trustworthy?" in *EGU General Assembly Conference Abstracts*, 2020, p. 13071.

[20] M. Galkin, J. Wu, E. Denis, and W. L. Hamilton, "Nodepiece: Compositional and parameter-efficient representations of large knowledge graphs," *arXiv preprint arXiv: 2106.12144*, 2021.

[21] A. Lacoste, A. Luccioni, V. Schmidt, and T. Dandres, "Quantifying the carbon emissions of machine learning," *arXiv preprint arXiv:1910.09700*, 2019.

[22] C. Burnicki. (2022) Cloud computing and carbon footprint. [Online]. Available: https://www.innoq.com/en/blog/cloud-computing-and-carbon-footprint/

[23] P. Henderson, J. Hu, J. Romoff, E. Brunskill, D. Jurafsky, and J. Pineau, "Towards the systematic reporting of the energy and carbon footprints of machine learning," *Journal of Machine Learning Research*, vol. 21, no. 248, pp. 1–43, 2020.

[24] A. Saltelli and S. Funtowicz, "When all models are wrong," *Issues in Science and Technology*, vol. 30, no. 2, pp. 79–85, 2014.

[25] Reddit.com. (2022) Data without context is noise! (with zoom). [Online]. Available: https://www.reddit.com/r/datascience/comments/tq93vt/data_without_context_is_noise_with_zoom/

[26] A. Zaveri, A. Rula, A. Maurino, R. Pietrobon, J. Lehmann, and S. Auer, "Quality assessment for linked data: A survey," *Semantic Web*, vol. 7, no. 1, pp. 63–93, 2016.

[27] T. Mitchell, W. Cohen, E. Hruschka, P. Talukdar, B. Yang, J. Betteridge, A. Carlson, B. Dalvi, M. Gardner, B. Kisiel *et al.*, "Never-ending learning," *Communications of the ACM*, vol. 61, no. 5, pp. 103–115, 2018.

[28] P. B. De Laat, "Algorithmic decision-making based on machine learning from big data: can transparency restore accountability?" *Philosophy & technology*, vol. 31, no. 4, pp. 525–541, 2018.

[29] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne *et al.*, "The fair guiding principles for scientific data management and stewardship," *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.

[30] D. Zheng, X. Song, C. Ma, Z. Tan, Z. Ye, J. Dong, H. Xiong, Z. Zhang, and G. Karypis, "Dgl-ke: Training knowledge graph embeddings at scale," in *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020, pp. 739–748.

[31] A. Lerer, L. Wu, J. Shen, T. Lacroix, L. Wehrstedt, A. Bose, and A. Peysakhovich, "Pytorch-biggraph: A large-scale graph embedding system," *arXiv preprint arXiv:1903.12287*, 2019.

[32] F. Lecue, "On the role of knowledge graphs in explainable ai," *Semantic Web*, vol. 11, no. 1, pp. 41–51, 2020.

[33] M. Palmonari and P. Minervini, "Knowledge graph embeddings and explainable ai," *Knowledge Graphs for Explainable Artificial Intelligence: Foundations, Applications and Challenges, IOS Press,, Amsterdam*, pp. 49–72, 2020.

[34] F. Bakhshandegan Moghaddam, C. Draschner, J. Lehmann, and H. Jabeen, "Literal2feature: An automatic scalable rdf graph feature extractor," in *Further with Knowledge Graphs*. IOS Press, 2021, pp. 74–88.

[35] N. Frosst and G. Hinton, "Distilling a neural network into a soft decision tree," *arXiv preprint arXiv:1711.09784*, 2017.

[36] I. Tiddi and S. Schlobach, "Knowledge graphs as tools for explainable machine learning: A survey," *Artificial Intelligence*, vol. 302, p. 103627, 2022.

[37] J. Urbani, S. Dutta, S. Gurajada, and G. Weikum, "Kognac: efficient encoding of large knowledge graphs," *arXiv preprint arXiv:1604.04795*, 2016.

[38] M. Sachan, "Knowledge graph embedding compression," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, pp. 2681–2691.

[39] C. F. Draschner, C. Stadler, F. Bakhshandegan Moghaddam, J. Lehmann, and H. Jabeen, "Distrdf2ml-scalable distributed in-memory machine learning pipelines for rdf knowledge graphs," in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 2021, pp. 4465–4474.

[40] N. Shadowen, "Ethics and bias in machine learning: A technical study of what makes us "good"," in *The Transhumanism Handbook*. Springer, 2019, pp. 247–261.

[41] N. M. Barbosa and M. Chen, "Rehumanized crowdsourcing: a labeling framework addressing bias and ethics in machine learning," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[42] BBC.com. (2022) Google apologises for photos app's racist blunder. [Online]. Available: https://www.oed.com/view/Entry/111850?redirectedFrom=machine+learning+#eid38479194

[43] T. Verge. (2022) Twitter taught microsoft's ai chatbot to be a racist asshole in less than a day. [Online]. Available: https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist

[44] technologyreview.com. (2022) Training a single ai model can emit as much carbon as five cars in their lifetimes. [Online]. Available: https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/

[45] E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in nlp," *arXiv preprint arXiv: 1906.02243*, 2019.

[46] M. Albooyeh, R. Goel, and S. M. Kazemi, "Out-of-sample representation learning for multi-relational graphs," *arXiv preprint arXiv:2004.13230*, 2020.

[47] D. Esteves, D. Moussallem, C. B. Neto, T. Soru, R. Usbeck, M. Ackermann, and J. Lehmann, "Mex vocabulary: a lightweight interchange format for machine learning experiments," in *Proceedings of the 11th International Conference on Semantic Systems*, 2015, pp. 169–176.

[48] G. C. Publio, D. Esteves, A. Ławrynowicz, P. Panov, L. Soldatova, T. Soru, J. Vanschoren, and H. Zafar, "Ml-schema: exposing the semantics of machine learning with schemas and ontologies," *arXiv preprint arXiv:1807.05351*, 2018.