Knowledge Graphs and Explainable AI in Healthcare

Enayat Rajabi [1] and **Somayeh Kafaie** [2,*]

1. Department of Financial and Information Management, Cape Breton University, Sydney, NS B1P 6L2, Canada
2. Mathematics and Computing Science Department, Saint Mary's University, 912 Robie Street, Halifax, NS B3H 3C3, Canada
* Correspondence: somayeh.kafaie@smu.ca

Abstract: Building trust and transparency in healthcare can be achieved using eXplainable Artificial Intelligence (XAI), as it facilitates the decision-making process for healthcare professionals. Knowledge graphs can be used in XAI for explainability by structuring information, extracting features and relations, and performing reasoning. This paper highlights the role of knowledge graphs in XAI models in healthcare, considering a state-of-the-art review. Based on our review, knowledge graphs have been used for explainability to detect healthcare misinformation, adverse drug reactions, drug-drug interactions and to reduce the knowledge gap between healthcare experts and AI-based models. We also discuss how to leverage knowledge graphs in pre-model, in-model, and post-model XAI models in healthcare to make them more explainable.

Keywords: eXplainable AI; healthcare; knowledge graph; explainability

1. Introduction

Artificial Intelligence (AI) systems have facilitated the automation and feasibility of many complicated tasks previously done manually by medical experts [1]. However, the lack of transparency of the complex models limits the understandability of healthcare practitioners and professionals. Designing an eXplainable AI (XAI) system provides an overview of an AI system, educates users, and helps in their future explorations. It makes AI systems more understandable, interpretable, and responsible [2]. By creating more explainable models, humans can better understand an AI system and its decisions [3]. It also enhances the trust of healthcare professionals and facilitates their decision-making process. Hence, the conclusions derived from data can be understandable for medical doctors. Recent XAI approaches assist AI adoption in clinical settings by building trust and transparency in traditional models.

Explainability can be understood as a characteristic of an AI-driven system allowing a person to reconstruct why a specific AI came up with its predictions [4]. It can usually be classified as global or local, based on the level at which explanations are provided [5]. Global explanations offer an understanding of the entire model's behaviour and reasoning, leading to expected outcomes in general. Local explanations facilitate the reasons for a single prediction to defend why the model made a specific decision for any particular instance. Combinations of both are also an interesting area of exploration. Another classification of explainability is in terms of applicability. Explanation techniques can be model agnostic, i.e., applied to any machine learning algorithms, or model-specific, which are applicable only for a single type or class of algorithm [5]. Explainability methods can also be classified into ante hoc and post hoc explainability methods. In the case of ante hoc methods, explainability is intrinsic or explainable by the model design itself. They are also referred to as transparent or white-box approaches. Post hoc methods are typically model agnostic and do not necessarily explain how black-box models work, but they may provide local explanations for a specific decision [6].

On the other hand, knowledge graphs have recently been used in healthcare to structure information, extract features and relationships and provide reasoning using ontologies [7]. A knowledge graph is a data structure representing entities as vertices and their relationships as directed labelled edges, which integrates and manipulates large-scale data from diverse sources [8]. Information in knowledge graphs can be organized in a hierarchical or graph structure in a way to allow an AI-based system to perform reasoning over the graph and answer more sophisticated queries ('questions') in some meaningful way [9,10]. Integrating heterogeneous information sources in the form of knowledge graphs allows healthcare systems to be intelligent to infer new facts and concepts. Knowledge graphs can provide human-understandable explanations, be integrated into the model, and add valuable additional knowledge. This article reviews the state-of-the-art studies that leverage knowledge graphs to explain the AI-based systems in healthcare and discusses the role and benefits of knowledge graphs in XAI to healthcare data and application.

2. State-of-the-Art

Knowledge graphs have been investigated in different areas of healthcare, including prioritizing cancer genes [11], identifying proteins' functions [12], drug repurposing [13,14], recognizing adverse drug reactions [15,16], drug–drug reactions [17], finding safer drugs [18,19] and detecting healthcare misinformation [20,21]. Some studies have created knowledge graphs to model medical information in graphs. In a review paper, Zeng et al. [22] summarized commonly used databases for knowledge graph construction and presented an overview of representative knowledge embedding models and knowledge graph-based predictions in the drug discovery field. Their study highlighted that the drug discovery process could be accelerated using knowledge graphs to assist data-driven pharmaceutical research.

Predicting drug–drug interactions and unknown adverse drug reactions (ADR) for new drugs are among the topics for which knowledge graphs show promising results. Knowledge graphs' ability to integrate diverse and heterogeneous sources of information has made them a common choice for better drug–drug interaction prediction and ADR detection [22]. Lin et al. proposed an end-to-end drug–drug interactions (DDI) framework to identify the correlations between drugs and other entities [17]. They presented a Knowledge Graph Neural Network (KGNN) model to resolve the DDI prediction. Also, Wang et al. applied machine learning methods to construct a knowledge graph with tumour, biomarker, drug and ADR as the type of its nodes [15]. The constructed knowledge graph has been used not only to discover potential ADRs of antitumor medicines, but also for explanations by providing the paths of "tumor-biomarker-drug" in the graph. In the constructed knowledge graph based on literature data, vertices represent the entities of the four types. There are undirected weighted edges where the weight shows the correlation (distance) between two vertices. The correlations were calculated using a naive Bayesian model [23] and considering the frequency of co-occurrences of two entities in the database. A correlation above a certain threshold indicates an edge between the corresponding entities. All pairs of drugs and their corresponding ADRs were collected for ADR discovery. Then, depth-first searches were used to find every path between the drug and ADR in the graph, which can also explain new ADR discoveries.

As another example of ADR, Bresso et al. investigated adverse drug reactions' molecular mechanisms by presenting models to distinguish between causative drugs or not [16]. Their proposed methodology is based on a graph-based feature construction method. They created a knowledge graph for pharmacogenomics (PGx), filtered out the noisy features, and identified three elements of drugs: paths, path patterns, and neighbours. They isolated both predictive and interpretative features, hypothesizing that they are explanatory for the classification and ADR mechanisms. They used decision trees and propositional rule learners over the extracted features to provide human-readable and explainable models. In another study [14], authors addressed the opacity intrinsic to the mathematical concepts that limit the use of AI in drug re-purposing (DR), i.e., using an approved drug to treat

a different disease, in the pharmaceutical industry. An end-to-end framework based on a knowledge-based computational DR model was proposed in their work. The authors built a knowledge graph (OREGANO) from various free and open-source knowledge bases. They used a reinforcement learning model to perform and describe possible paths and links between entities, which were drugs and diseases, in this study.

Teng et al. addressed the time-consuming, error-prone, and expensive International Classification of Diseases (ICD) coding process using feature extraction in knowledge graphs [24]. Most existing ICD coding models only translate the simple diagnosis descriptions into ICD codes, obscuring the reasons and details behind specific diagnoses. Their proposed approach presents automatic ICD code assignment using clinical records. Knowledge graphs and attention mechanisms were expanded into medical code prediction to improve the explainability of medical codes. They used a Multiple Convolutional Neural Networks (Multi-CNN) model to capture local correlation and extract the key features from clinical records. They also leveraged Structural Deep Network Embedding (SDNE) to enrich the meaning of terminologies through integrated related knowledge points for the ICD using the Freebase database. They used an attention mechanism to help understand the meaning of associated terminologies and make the coding results interpretable.

Furthermore, knowledge graph techniques can be incorporated into identifying drug–disease association and drug–target interaction, which are major steps in drug repurposing. Recently, there have been interesting studies on COVID-19 treatment based on drug repurposing, which construct and use knowledge graphs like DRKG [25] and COVID-KG [13] for effective extraction of semantic information and relations. For a list of conducted research in this area, we refer readers to [22].

Regarding detecting misinformation in healthcare, ref. [21] presents an explainable COVID-19 misinformation detection model called HC-COVID to address the critical problem of COVID-19 misinformation. HC-COVID is a hierarchical crowd-sourced knowledge graph-based framework that explicitly models the COVID-19 knowledge facts to accurately identify the related knowledge facts and explain the detection results. After constructing a crowd-source hierarchical knowledge graph (CHKG), they identified COVID-19 specific knowledge facts by developing a multi-relational graph neural network to encode input COVID-19 claims and integrate the claim information with COVID-19 specific knowledge facts in CHKG. The attention outputs are used to retrieve informative graph triples from CHKG as explanations for the COVID-19 misinformation detection results. HC-COVID explains why the COVID-19 claim is misleading based on the correlation between the COVID-19 claim and the corresponding knowledge facts in the knowledge graph. DETERRENT [20] is another XAI study based on a knowledge graph for detecting healthcare misinformation that uses knowledge graphs and attention networks to explain misinformation detection in social contexts. This study shows that knowledge graphs can provide valuable explanations for the results of information detection. It has an information propagation net to propagate the knowledge between articles and nodes, knowledge-aware attention to learn the weights of a node's neighbours in the knowledge graph and aggregates the information from the neighbours and an article's contextual information to update its representation. It also has a prediction layer to take an article's representation and output a predicted label. Furthermore, they use a Bayesian personalized ranking (BPR) loss to capture both positive and negative relations in the knowledge graph (e.g, Heals vs DoesNotHeal).

Knowledge graphs are also used to minimize the knowledge gap between domain experts and AI-based models. Díaz-Rodríguez et al. propose an eXplainable Neural-symbolic learning (X-NeSyL) methodology [26] to learn both symbolic and deep representations, together with an explainability metric to assess the level of alignment of machine and human expert explanations. It uses an extended knowledge representation component, in the form of a knowledge graph, with a neural representation learning model to compare the machine and the expert knowledge. This method uses a knowledge graph based on the Web Ontology Language (OWL) (https://www.w3.org/TR/owl2-overview/, accessed on

6 September 2022). It presents a new training procedure based on feature attribution to enhance the interpretability of the classification model. GNN-SubNet is another graph-based deep learning framework for disease subnetwork detection via explainable graph neural networks (GNN) [27]. By integrating a knowledge graph into the algorithmic pipeline, their proposed model allows the expert-in-the-loop and ultimately provides accurate predictions with explanations. Each patient in their system is represented by the topology of a protein–protein network (PPI). The nodes are enriched by multimodal molecular data, such as gene expression and DNA methylation. They leveraged GNN graph classification to classify patients into specific and randomized groups. The decisions of the GNN classifier are fed into a GNN explainer algorithm (a post-hoc method) to obtain the node importance values from which edge relevant scores are computed for explainability. GNN-SubNet is capable of reporting on both positive contributions and the negative contribution of features to a particular prediction in graphical representation.

In Table 1, we summarize the literature review by categorizing the reviewed articles based on how they leveraged knowledge graphs in XAI models in healthcare.

Table 1. Knowledge graph applications in Healthcare XAI in our literature review.

Knowledge Graph Application	Article(s)
Detecting healthcare misinformation	[20,21]
Adverse drug reactions	[15,16]
Drug re-purposing	[14]
Feature extraction for disease coding process	[24]
Drug-disease association and drug-target interaction	[13,22,25]
Minimizing the knowledge gap between healthcare experts and AI-based models	[26]
Drug-drug interactions	[17,22]

3. The Role of Knowledge Graphs in XAI and Healthcare

As mentioned earlier, knowledge graphs are applied in different ways in healthcare to increase the explainability of AI models. They annotate, organize, and present different types of information meaningfully and add semantic labels to healthcare datasets. The graph-based structure of knowledge graphs based on subject–predicate–object is used to represent concepts and their relationships in an XAI model [28]. These relation triples, for example, in fact-checking, can change the credibility of specific medical information and explain the detection results [20,29,30].

Knowledge graphs can visualize data such that related information is constantly connected [31] and can help obtain in-depth insights into the underlying mechanisms. Therefore, they can structure the conceptual information to be more understandable to humans. Knowledge graphs can also encode contexts, expose connections and relations, and support inference and causation natively [7]. While adding new data in many traditional big data schemes is challenging and requires understanding the entire framework, updates in knowledge graphs are straightforward and much easier [31].

In the following subsections, we will describe the role of knowledge graphs in XAI models in healthcare from different aspects, such as knowledge graphs for healthcare data, knowledge graph construction, and the application of knowledge graphs in healthcare.

3.1. Knowledge Graphs and Health Data

We can divide health care data into three categories: omics data, clinical data, and sensor data. Omics data deal with large-scale and high-dimensional data, which indicates the totality of some kind like genomics, proteomics and transcriptomics [32]. Clinical data contains patients' electronic health records, which are collected as part of ongoing patient care and act as the main resource for most medical and health research [33]. On the other hand, sensor data refers to data collected from numerous wearable and wireless sensor

devices such as pulse oximetry, heart rate monitors, blood pressure cuffs, Parkinson's disease monitoring system, and depression-mood monitoring systems [34].

Knowledge graphs constructed based on omic data, more specifically multi-omic data, are used to explore the relationship between such entities and diseases and provide novel discoveries [35]. These graphs have been used to identify protein–protein interactions [36–39], miRNA–disease associations [40] and gene–disease association [41–43]. Knowledge graphs can be used to extract novel information from clinical data and improve patient care. Studies have been conducted to recommend safer medicine combinations [44,45], predict the probability of patient–disease associations like heart failure probability [46] and improve patient diagnoses [45,47].

Knowledge graphs provide the opportunity to convert a variety of healthcare data into a uniform graph format and picture all the links among different objects to create knowledge from diverse fields, like diseases, drugs or treatments, through edges with various labels. This ability of knowledge integration, not available in traditional pharmacologic experiments, can speed up the discovery of knowledge [15].

3.2. Knowledge Graph Construction in XAI Healthcare

A knowledge graph embodies human understanding by finding inherent relations between concepts and entities and providing links between them and forms a conceptual graph network that not only improves the overall explainability of complex AI algorithms [48], but can also improve detection performance [20]. These graphs can be categorized as "common-sense" (i.e., data on everyday world), "factual" (i.e., knowledge about facts and events) and "domain" (i.e., knowledge from a certain domain) [8]. However, in healthcare, domain knowledge graphs are usually constructed and applied.

Knowledge graphs are constructed from various sources of unstructured databases and structured ontologies [22], like medical images, unstructured text and reports, electronic medical records (EMRs), etc. Graph entities and their relations can also be extracted from published biomedical literature applying text mining technologies [22,35]. The examples include GNBR [49] and DRKG [25].

Traditionally, biomedical knowledge graphs are constructed from manually curated databases by domain experts [35]. However, in recent years, automatic methods relying on machine learning and natural language processing techniques have become common tools for graph construction. Nicholson et al. group them into three categories of rule-based extraction, unsupervised machine learning, and supervised machine learning [35]. Although automatic approaches are significantly faster, as one can expect, studies show that they are unreliable in dealing with ambiguity embedded in human language and extracting less-common relations [50].

3.3. Knowledge Graphs Application In Healthcare XAI

There have been many attempts in the literature to provide transparent and understandable AI models with explainable and trustworthy outcomes, especially in recent years. In this section, we categorize applications of knowledge graphs in the healthcare domain toward improving the explainability of AI models.

- Entity/relation extraction: Narrative of patients' interactions are usually provided in clinical notes in the form of unstructured data and free text in healthcare. This information is transformed into a structured format such as a named entity or common vocabulary. Knowledge graphs are used to represent clinical notes with named entity recognition methods and map them to vocabularies using named entity normalization techniques. They are also used in relation extraction, where the semantic relation is typically extracted between two entities [51]. For example, in a disease knowledge graph, the relationship between disease and other concepts, such as diagnosis or treatment, can be extracted using different relation extraction algorithms.
- Enrichment: Knowledge graphs are employed to enrich datasets with internal or external information and knowledge. The ability to provide a trustable and explainable

machine learning model with high prediction accuracy can be improved if the AI system is enriched by additional knowledge.

- Inference and reasoning: Knowledge graphs usually leverage deduction reasoning to help infer new facts and knowledge. Reasoning over a knowledge graph is an evidence-based approach that is more acceptable and interpretable for clinicians. For example, EHR data can be transformed into a semantic net model (patient-centralized) under a knowledge graph to create an EHR data trajectory and reasoning using semantic rules. Designing such a system allows reasoning to identify critical clinical discoveries within EHR data and presents the clinical significance for clinicians to understand the neglected information better [52].
- Explanations and visualizations: The XAI models provide explanations for physicians and healthcare professionals so that the outputs are understandable and transparent. The knowledge graphs help provide more insights into the reasons for model predictions and can also represent the results in graphs. Human-in-the-loop techniques can also be used to validate the results or refine the knowledge graph to achieve high accuracy and better explainability.

4. Discussion

XAI and knowledge graphs can improve each other from different aspects. Knowledge graphs can be used in different sections of an AI-based healthcare system. As there are various sources of information in healthcare (e.g., texts with different formats, images, and sensors) with diverse data structures, integrating them all into a single graph is a complicated task. As discussed in Section 3.2, machine learning techniques can be applied to automate and facilitate knowledge graph construction from diverse types of healthcare data. However, it is challenging to define a graph structure for various data types and represent the required data in the least complex format, which can be easily and with low time complexity accessible. In terms of machine learning techniques, Natural Language Processing (NLP) has been mostly used in the literature to extract entities and their relations from medical text. However, they might add noise to the graph and reduce its accuracy by introducing entities with wrong names or invalid relations [22,53]. In fact, quality estimation of knowledge graphs is still a challenging task. As a rare example, Zhao et al. [54] applied logic rules to estimate the probability of knowledge graph triplets. However, more studies are required to verify the quality of knowledge graphs by designing feasible methods [22].

In terms of leveraging knowledge graphs application inside of XAI models, they play a central role in the new design of deep learning models by adding logic representation layers, encoding the semantics of inputs, outputs and their properties for causation and feature reasoning [7]. For example, to predict human behaviours in healthy social networks, Phan et al. [55] proposed an ontology-based deep learning model as a Restricted Boltzmann machine where it can also provide explanations for the predicted behaviour as a set of triples that maximizes the likelihood of a behaviour. On the other hand, the information embedded in knowledge graphs can be used to enhance the result of XAI models and explain or adjust the model's output. However, the quality of the explanations largely depends on the precision of the knowledge graph and its construction, and some validation methods might be required to verify them.

To better understand the application of knowledge graphs in healthcare XAI, we categorized their roles and applications according to the literature review. Figure 1 illustrates that various healthcare data can be leveraged to construct a knowledge graph. The knowledge graph can be created after extracting features or relations in a text or image. It can be infused with additional knowledge from the Web of Data or some semantic similarities techniques. Knowledge graphs can be used to enrich the training datasets in machine learning models and can be used for reasoning and querying the data. The combination of graph representation with AI models can be applied to make predictions within genomic, pharmaceutical, and clinical domains [35]. Eventually, knowledge graphs

can be reviewed by domain experts (such as healthcare professionals or physicians) and applied again in the healthcare system to improve the performance of machine learning models or prediction accuracy.
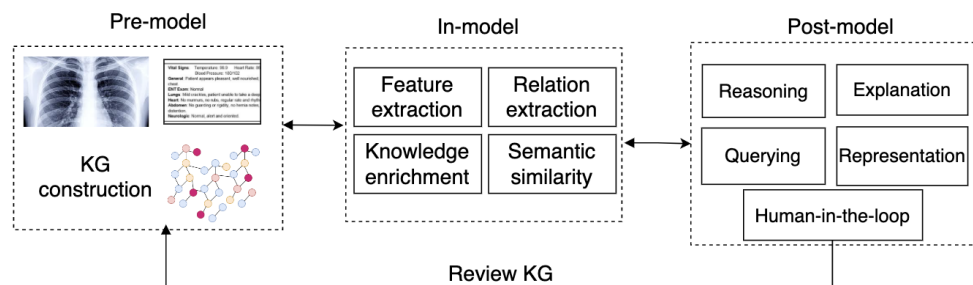
Figure 1. The figure shows how an XAI model can use knowledge graphs for explainability.

On the other side, XAI methods can be used to discover new knowledge from knowledge graphs by grouping nodes, link prediction or node classification [56]. Graph representation learning [57] approaches, as an example, can be used in healthcare to encode the network structure into low-dimensional space of dense vectors that often are assigned to nodes [58], but can also embed edges [59].

When dealing with substantial knowledge graphs with many vertices and a high outdegree (i.e., many links from a given entity to the others), graph traversal in the knowledge discovery process is a significant concern. Identifying the most relevant paths and closest facts among many available ones in such knowledge graphs might be challenging. Another challenge might be pruning the graph to handle noise and filtering irrelevant entities. Furthermore, the application of knowledge graphs for extracting facts and explanations can be extended to more than one-hop neighbourhood of vertices. In fact, more valuable knowledge might be retrieved by exploring nodes' neighbourhoods in depth and providing a chain of facts; a task that can be compute-intensive, especially for giant knowledge graphs.

Despite the advantages of XAI models described in this paper, many studies suggest that XAI can still not fulfil its intended mission [7,60,61]. Flat representation of data without appropriate context considerations is one of the main challenges of the existing XAI approaches that can, at least partially, be addressed by knowledge graphs by providing better data representation and more interpretable models [7]. As discussed by De bruijn et al. [61], XAI mainly focuses on providing explanations understandable by the public; however, it is challenging, as many people do not have enough expertise to assess the quality of AI decisions. Knowledge graphs can help alleviate this issue by providing visualizations and semantic representations of concepts. In healthcare, the considerable diversity of data and its complexity might create a trade-off between explainability and performance of an XAI model [60]. Furthermore, the model complexity can cause learning biases in the model for certain types of biomedical data, which affect the quality of results and their explainability. In critical applications like healthcare, fixing these learning flaws can have a higher priority than XAI [60]. All in all, although XAI is a promising and essential research direction in healthcare, it is still in its infancy, and further investigation is required to overcome its challenges.

5. Conclusions

Knowledge graphs have been widely utilized to explain drug–drug interactions, identify misinformation in clinical settings, reduce human knowledge and machine gaps, explain clinical notes and prescriptions, and enrich healthcare data with additional knowledge. Combining knowledge graphs with machine learning models provides more insights into making AI-based models more explainable. This paper categorized the knowledge graphs' applications in the healthcare domain after a state-of-the-art review. We demonstrated how knowledge graphs are used in healthcare systems for explainability purposes in different

studies for entity and relation extraction, knowledge graph construction, reasoning, and knowledge representation.

Author Contributions: Conceptualization, E.R. and S.K.; methodology, E.R. and S.K.; literature review and investigation, E.R. and S.K.; writing—original draft preparation, E.R. and S.K.; writing—review and editing, E.R. and S.K.; visualization, E.R. and S.K.; project administration, E.R.; funding acquisition, E.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been partially funded by NSERC (Natural Sciences and Engineering Research Council) Discovery Grant (RGPIN-2020-05869) in Canada.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Huang, J.; Shlobin, N.A.; Lam, S.K.; DeCuypere, M. Artificial intelligence applications in pediatric brain tumor imaging: A systematic review. *World Neurosurg.* **2022**, *157*, 99–105. [CrossRef]
2. Wohlin, C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, London, UK, 13–14 May 2014; pp. 1–10.
3. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]
4. Amann, J.; Blasimme, A.; Vayena, E.; Frey, D.; Madai, V.I. Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 310. [CrossRef]
5. Adadi, A.; Berrada, M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160. [CrossRef]
6. Antoniadi, A.M.; Du, Y.; Guendouz, Y.; Wei, L.; Mazo, C.; Becker, B.A.; Mooney, C. Current challenges and future opportunities for xai in machine learning-based clinical decision support systems: A systematic review. *Appl. Sci.* **2021**, *11*. [CrossRef]
7. Lecue, F. On the Role of Knowledge Graphs in Explainable AI. *Semant. Web* **2020**, *11*, 41–51. doi: 10.3233/SW-190374. [CrossRef]
8. Tiddi, I.; Schlobach, S. Knowledge Graphs as Tools for Explainable Machine Learning: A Survey. *Artif. Intell.* **2022**, *302*, 103627. [CrossRef]
9. Kejriwal, M. *Domain-Specific Knowledge Graph Construction*; Springer: Berlin/Heidelberg, Germany, 2019.
10. Telnov, V.; Korovin, Y. Semantic Web and Interactive Knowledge Graphs as an Educational Technology. In *Cloud Computing Security-Concepts and Practice*; IntechOpen: Londong, UK, 2020.
11. Shang, H.; Liu, Z.P. Network-based prioritization of cancer genes by integrative ranks from multi-omics data. *Comput. Biol. Med.* **2020**, *119*, 103692. [CrossRef]
12. Crichton, G.; Guo, Y.; Pyysalo, S.; Korhonen, A. Neural Networks for Link Prediction in Realistic Biomedical Graphs: A Multi-dimensional Evaluation of Graph Embedding-based Approaches. *BMC Bioinform.* **2018**, *19*, 176. [CrossRef]
13. Wang, Q.; Li, M.; Wang, X.; Parulian, N.; Han, G.; Ma, J.; Tu, J.; Lin, Y.; Zhang, R.H.; Liu, W.; et al. COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, 6–11 June 2021; pp. 66–77. [CrossRef]
14. Drancé, M.; Boudin, M.; Mougin, F.; Diallo, G. Neuro-symbolic XAI for Computational Drug Repurposing. In Proceedings of the 13th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2021)-Volume 2: KEOD, Online, 25–27 October 2021; pp. 220–225. [CrossRef]
15. Wang, M.; Ma, X.; Si, J.; Tang, H.; Wang, H.; Li, T.; Ouyang, W.; Gong, L.; Tang, Y.; He, X.; et al. Adverse Drug Reaction Discovery Using a Tumor-Biomarker Knowledge Graph. *Front. Genet.* **2021**, *11*, 1737. [CrossRef]
16. Bresso, E.; Monnin, P.; Bousquet, C.; Calvier, F.E.; Ndiaye, N.C.; Petitpain, N.; Smaïl-Tabbone, M.; Coulet, A. Investigating ADR mechanisms with Explainable AI: A feasibility study with knowledge graph mining. *BMC Med. Inform. Decis. Mak.* **2021**, *21*, 171–184. [CrossRef] [PubMed]
17. Lin, X.; Quan, Z.; Wang, Z.J.; Ma, T.; Zeng, X. KGNN: Knowledge Graph Neural Network for Drug-Drug Interaction Prediction. *IJCAI* **2020**, *380*, 2739–2745.
18. Shang, J.; Xiao, C.; Ma, T.; Li, H.; Sun, J. GAMENet: Graph Augmented Memory Networks for Recommending Medication Combination. In Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; AAAI'19/IAAI'19/EAAI'19. [CrossRef]
19. Gong, F.; Wang, M.; Wang, H.; Wang, S.; Liu, M. SMR: Medical Knowledge Graph Embedding for Safe Medicine Recommendation. *Big Data Res.* **2021**, *23*, 100174. [CrossRef]

20.   Cui, L.; Seo, H.; Tabar, M.; Ma, F.; Wang, S.; Lee, D. DETERRENT: Knowledge Guided Graph Attention Network for Detecting Healthcare Misinformation. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Anchorage, AK, USA, 4–8 August 2020; pp. 492–502. [CrossRef]

21.   Kou, Z.; Shang, L.; Zhang, Y.; Wang, D. Hc-Covid. *Proc. ACM Hum.-Comput. Interact.* **2022**, *6*, 36. [CrossRef]

22.   Zeng, X.; Tu, X.; Liu, Y.; Fu, X.; Su, Y. Toward Better Drug Discovery with Knowledge Graph. *Curr. Opin. Struct. Biol.* **2022**, *72*, 114–126. [CrossRef]

23.   Bayes' Theorem (Stanford Encyclopedia of Philosophy). Available online': https://plato.stanford.edu/entries/bayes-theorem/ (accessed on 22 September 2022).

24.   Teng, F.; Yang, W.; Chen, L.; Huang, L.F.; Xu, Q. Explainable Prediction of Medical Codes With Knowledge Graphs. *Front. Bioeng. Biotechnol.* **2020**, *8*, 867. [CrossRef]

25.   Zeng, X.; Song, X.; Ma, T.; Pan, X.; Zhou, Y.; Hou, Y.; Zhang, Z.; Li, K.; Karypis, G.; Cheng, F. Repurpose Open Data to Discover Therapeutics for COVID-19 Using Deep Learning. *J. Proteome Res.* **2020**, *19*, 4624–4636. [CrossRef]

26.   Díaz-Rodríguez, N.; Lamas, A.; Sanchez, J.; Franchi, G.; Donadello, I.; Tabik, S.; Filliat, D.; Cruz, P.; Montes, R.; Herrera, F. EXplainable Neural-Symbolic Learning (X-NeSyL) methodology to fuse deep learning representations with expert knowledge graphs: The MonuMAI cultural heritage use case. *Inf. Fusion* **2022**, *79*, 58–83. [CrossRef]

27.   Pfeifer, B.; Secic, A.; Saranti, A.; Holzinger, A. GNN-SubNet: Disease subnetwork detection with explainable Graph Neural Networks. *Bioinformatics* **2022**, *38*, ii120–ii126. [CrossRef]

28.   Li, L.; Wang, P.; Yan, J.; Wang, Y.; Li, S.; Jiang, J.; Sun, Z.; Tang, B.; Chang, T.-H.; Wang, S.; et al. Real-world data medical knowledge graph: Construction and applications. *Artif. Intell. Med.* **2020**, *103*, 101817. [CrossRef]

29.   Ciampaglia, G.L.; Shiralkar, P.; Rocha, L.M.; Bollen, J.; Menczer, F.; Flammini, A. Computational Fact Checking from Knowledge Networks. *PLoS ONE* **2015**, *10*, e0128193. [CrossRef]

30.   Huynh, V.P.; Papotti, P. A Benchmark for Fact Checking Algorithms Built on Knowledge Bases. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; CIKM '19, p. 689. [CrossRef]

31.   Admin. Knowledge Graphs: Backbone of Data-Driven Culture in Life Sciences. 2019. Available online: https://www.virtusa.com/perspectives/article/knowledge-graphs-backbone-of-data-driven-culture-in-life-sciences (accessed on 7 August 2022).

32.   Vailati-Riboni, M.; Palombo, V.; Loor, J.J., What Are Omics Sciences? In *Periparturient Diseases of Dairy Cows: A Systems Biology Approach*; Ametaj, B.N., Ed.; Springer International Publishing: Cham, Switzerland, 2017; pp. 1–7. [CrossRef]

33.   Maloy, C. Library Guides: Data Resources in the Health Sciences: Clinical Data. 2022. Available online: https://guides.lib.uw.edu/hsl/data/findclin (accessed on 16 July 2022).

34.   Sow, D.; Turaga, D.S.; Schmidt, M., Mining of Sensor Data in Healthcare: A Survey. In *Managing and Mining Sensor Data*; Aggarwal, C.C., Ed.; Springer: Boston, MA, USA, 2013; pp. 459–504. [CrossRef]

35.   Nicholson, D.N.; Greene, C.S. Constructing Knowledge Graphs and Their Biomedical Applications. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1414–1428. [CrossRef] [PubMed]

36.   Wang, H.; Huang, H.; Ding, C.; Nie, F. Predicting Protein-Protein Interactions from Multimodal Biological Data Sources via Nonnegative Matrix Tri-Factorization. In *Proceedings of Research in Computational Molecular Biology*; Chor, B., Ed.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 314–325.

37.   Wu, Q.; Wang, Z.; Li, C.; Ye, Y.; Li, Y.; Sun, N. Protein Functional Properties Prediction in Sparsely-label PPI Networks through Regularized Non-negative Matrix Factorization. *BMC Syst. Biol.* **2015**, *9*, S9. [CrossRef]

38.   Alshahrani, M.; Khan, M.A.; Maddouri, O.; Kinjo, A.R.; Queralt-Rosinach, N.; Hoehndorf, R. Neuro-symbolic Representation Learning on Biological Knowledge Graphs. *Bioinformatics* **2017**, *33*, 2723–2730. [CrossRef]

39.   Trivodaliev, K.; Josifoski, M.; Kalajdziski, S. Deep Learning the Protein Function in Protein Interaction Networks. In Proceedings of the ICT Innovations 2018. Engineering and Life Sciences, Ohrid, Macedonia, 17–19 September 2018; Kalajdziski, S.; Ackovska, N., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 185–197.

40.   Wu, F.X.; Shen, Z.; Zhang, Y.H.; Han, K.; Nandi, A.K.; Honig, B.; Huang, D.S. miRNA-Disease Association Prediction with Collaborative Matrix Factorization. *Complexity* **2017**, *2017*, 2498957. [CrossRef]

41.   Yang, K.; Wang, N.; Liu, G.; Wang, R.; Yu, J.; Zhang, R.; Chen, J.; Zhou, X. Heterogeneous Network Embedding for Identifying Symptom Candidate Genes. *J. Am. Med Inform. Assoc.* **2018**, *25*, 1452–1459. [CrossRef] [PubMed]

42.   Xu, B.; Liu, Y.; Yu, S.; Wang, L.; Dong, J.; Lin, H.; Yang, Z.; Wang, J.; Xia, F. A Network Embedding Model for Pathogenic Genes Prediction by Multi-path Random Walking on Heterogeneous Network. *BMC Med. Genom.* **2019**, *12*, 188. [CrossRef] [PubMed]

43.   Wang, X.; Gong, Y.; Yi, J.; Zhang, W. Predicting Gene-disease Associations from the Heterogeneous Network using Graph Embedding. In Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 18–21 November 2019; pp. 504–511. [CrossRef]

44.   Wang, M.; Liu, M.; Liu, J.; Wang, S.; Long, G.; Qian, B. Safe Medicine Recommendation via Medical Knowledge Graph Embedding. *arXiv* **2017**, arXiv:1710.05980.

45.   Zhao, C.; Jiang, J.; Guan, Y.; Guo, X.; Hen, B. EMR-based Medical Knowledge Representation and Inference via Markov Random Fields and Distributed Representation Learning. *Artif. Intell. Med.* **2018**, *87*, 49–59. [CrossRef]

46. Choi, E.; Bahadori, M.T.; Song, L.; Stewart, W.F.; Sun, J. GRAM: Graph-based Attention Model for Healthcare Representation Learning. In Proceedings of the International Conference on Knowledge Discovery & Data Mining, Halifax, NS, Canada, 13–17 August 2017; pp. 787–795.

47. Wang, S.; Chang, X.; Li, X.; Long, G.; Yao, L.; Sheng, Q.Z. Diagnosis Code Assignment Using Sparsity-Based Disease Correlation Embedding. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 3191–3202. [CrossRef]

48. Sarker, M.K. Towards Explainable Artificial Intelligence (XAI) Based on Contextualizing Data with Knowledge Graphs. Ph.D. Thesis, Kansas State University, Manhattan, KS, USA, 2020.

49. Percha, B.; Altman, R.B. A Global Network of Biomedical Relationships Derived from Text. *Bioinformatics* **2018**, *34*, 2614–2624. [CrossRef] [PubMed]

50. Névéol, A.; Islamaj Doǧan, R.; Lu, Z. Semi-automatic Semantic Annotation of PubMed Queries: A Study on Quality, Efficiency, Satisfaction. *J. Biomed. Inform.* **2011**, *44*, 310–318. [CrossRef]

51. Haq, H.U.; Kocaman, V.; Talby, D. Deeper Clinical Document Understanding Using Relation Extraction. *arXiv* **2021**, arXiv:2112.13259.

52. Shang, Y.; Tian, Y.; Zhou, M.; Zhou, T.; Lyu, K.; Wang, Z.; Xin, R.; Liang, T.; Zhu, S.; Li, J. EHR-Oriented Knowledge Graph System: Toward Efficient Utilization of Non-Used Information Buried in Routine Clinical Practice. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 2463–2475. [CrossRef]

53. Song, B.; Li, F.; Liu, Y.; Zeng, X. Deep Learning Methods for Biomedical Named Entity Recognition: A Survey and Qualitative Comparison. *Briefings Bioinform.* **2021**, *22*, bbab282. [CrossRef] [PubMed]

54. Zhao, S.; Qin, B.; Liu, T.; Wang, F. Biomedical Knowledge Graph Refinement with Embedding and Logic Rules. *arXiv* **2020**, arXiv:2012.01031.

55. Phan, N.; Dou, D.; Wang, H.; Kil, D.; Piniewski, B. Ontology-based Deep Learning for Human Behavior Prediction with Explanations in Health Social Networks. *Inf. Sci.* **2017**, *384*, 298–313. [CrossRef]

56. Cassiman, J. How Are Knowledge Graphs and Machine Learning Related? 2022. Available online: https://blog.ml6.eu/how-are-knowledge-graphs-and-machine-learning-related-ff6f5c1760b5 (accessed on 26 August 2022).

57. Hamilton, W.L.; Ying, R.; Leskovec, J. Representation Learning on Graphs: Methods and Applications. *arXiv* **2017**, arXiv:1709.05584.

58. Grover, A.; Leskovec, J. node2vec: Scalable Feature Learning for Networks. In Proceedings of the International Conference on Knowledge Discovery & Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 855–864. [CrossRef]

59. Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; Yakhnenko, O. Translating Embeddings for Modeling Multi-relational Data. In Proceedings of Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 5–10 December 2013; Burges, C.; Bottou, L.; Welling, M.; Ghahramani, Z.; Weinberger, K., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2013; Volume 26.

60. Han, H.; Liu, X. The Challenges of Explainable AI in Biomedical Data Science. *BMC Bioinform.* **2022**, *22*, 443. [CrossRef]

61. de Bruijn, H.; Warnier, M.; Janssen, M. The Perils and Pitfalls of Explainable AI: Strategies for Explaining Algorithmic Decision-Making. *Gov. Inf. Q.* **2022**, *39*, 101666. [CrossRef]