IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 April 2024, accepted 24 June 2024, date of publication 3 July 2024, date of current version 12 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3422416

SURVEY

A Review of Explainable Recommender Systems Utilizing Knowledge Graphs and Reinforcement Learning

NEERAJ TIWARY[1], SHAHRUL AZMAN MOHD NOAH[1], (Member, IEEE), FARIZA FAUZI[2], AND TAN SIOK YEE[1]

[1]Centre for Artificial Intelligence Technology (CAIT), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor 43600, Malaysia
[2]Centre for Cyber Security, Faculty of Information Science and Technology, UKM, Bangi, Selangor 43600, Malaysia

Corresponding author: Shahrul Azman Mohd Noah (shahrul@ukm.edu.my)

This work was supported by Universiti Kebangsaan Malaysia through the Prime Impact Fund under Grant DIP-2020-017.

ABSTRACT This review paper addresses the research question of the significance of explainability in AI and the role of integrating KG and RL to enhance Explainable Recommender Systems (XRS). It surveys articles published from January 2015 to March 2024 on XRS, focusing on knowledge graphs (KGs) and reinforcement learning (RL) for achieving explainability in recommender systems. Employing a systematic methodology, it introduces a custom Python-based web scraper to efficiently navigate and extract relevant academic research papers from IEEE, ScienceDirect (Elsevier), ACM, and Springer online databases. The study encompasses the PRISMA methodology to conduct a thorough analysis and identify pertinent research works. This systematic literature review aims to provide a unified view of the field by reviewing eight existing XRS literature reviews and 29 pertinent XRS studies involving KG and RL from the specified period. It categorizes and analyses relevant research papers based on their implementation methodologies and explores significant contributions, encompassing perspectives on model-agnostic and model-intrinsic explanations.

INDEX TERMS Explainable artificial intelligence (XAI), explainable recommender systems (XRS), knowledge graph (KG), PRISMA, reinforcement learning (RL), systematic literature review (SLR), web scraper.

I. INTRODUCTION

In today's digital world, recommender systems (RSs) are pivotal in delivering personalized content recommendations across various domains. However, these systems often function as opaque black boxes, making it difficult for users to understand the reasoning behind receiving particular recommendations. Explainable Recommender Systems (XRS) address this transparency problem by providing accurate recommendations besides explaining decision-making processes to users [1], as depicted in Fig. 1.

This review paper aims to explore the XRS and their approaches, particularly utilizing knowledge graph (KG) and reinforcement learning (RL). The objective of the paper is to address the following specific research questions:

- RQ1: What contributions do existing survey or review papers make to the field of XRS when utilizing KG and RL?
- RQ2: What defines explainability in artificial intelligence (AI), and why is it important for RS?
- RQ3: How does KG contribute to achieving explainability within RS?
- RQ4: How does RL contribute to achieving explainability within RS?
- RQ5: How important is it to integrate KG and RL to enhance the effectiveness of the XRS?

Before initiating the detailed review process to address the aforementioned research questions, the paper surveyed the

The associate editor coordinating the review of this manuscript and approving it for publication was Mansoor Ahmed.

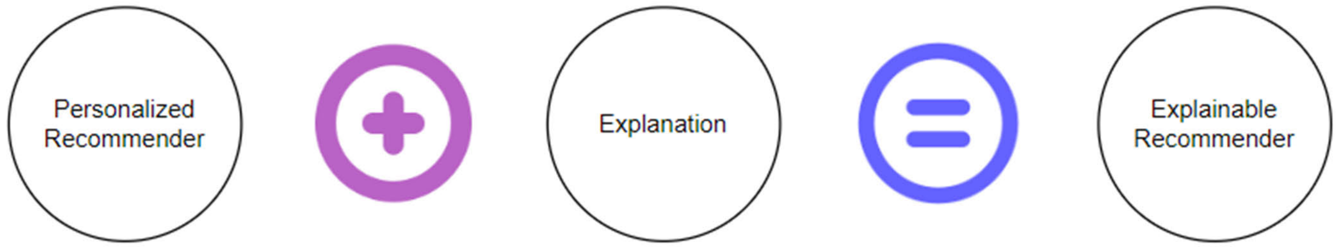

FIGURE 1. Explainable recommendations.

existing literature in the field, summarizing the findings in the Related Works section. The thorough analysis of prior studies underscored the need for a specific review paper focusing on XRS implementation employing both KG and RL.

This review paper provides an overview of RS and explainability in AI, KG, and RL, highlighting their roles in developing effective XRS. It systematically examines the advancements in XRS utilizing KG and RL by reviewing the pertinent research works from Jan 2015 to Mar 2024. Additionally, it introduces a custom web scraper tool designed to streamline the retrieval and organization of research papers from prominent platforms like IEEE, ScienceDirect, ACM, and Springer online databases. The web scraper tool, developed using Python, Microsoft EDGE web driver, and Selenium for automated web scraping, enhances access and organization of academic content. The effectiveness of the retrieved academic content relies on accurately crafted search queries that serve as the primary filter and meticulously analyze that in titles, abstracts, or keywords.

The review procedure follows the process outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology [2]. It conducts a thorough examination of selected papers, encompassing four sub-stages. Each sub-stage eliminates irrelevant articles based on specific criteria. The PRISMA approach ensures a curated collection of relevant research papers. It allows systematic quantitative and qualitative analyses of candidate research articles, summarising their findings, including chronological categorization of research papers, platform-wise distributions, and in-depth analyses of significant contributions.

Finally, the review paper delves into a detailed review and discussion of the prominent articles relevant to the objectives of this review paper. These articles provide a comprehensive view of the trends and advancements in the field of XRS utilizing KG and RL. This in-depth examination aims to extract valuable insights, highlight significant research trends, and contribute to a nuanced understanding of the state of XRS.

Thus, this review paper contributes with the following contributions:

- It offers an overview of existing reviews on XRS, summarizing their contributions and areas of focus.
- A tailored web scraping tool crafted to efficiently gather and organize research papers from leading platforms like IEEE, ScienceDirect, ACM, and Springer online databases.
- It conducts a comprehensive analysis and discussion of the research questions, exploring the need for explainability in RS and investigating implementation approaches with KG and RL.
- The review paper delves deeper into research advancements in XRS utilizing KG and RL by examining relevant research papers.
- The paper offers an overview of trends in XRS, recommends methods for XRS, and suggests future directions to address current challenges based on the studies reviewed.

The paper's structure is designed to offer a comprehensive exploration of XRS utilizing KG and RL. Section II examines related surveys and review papers, summarizing their contributions. Section III offers foundational knowledge on relevant concepts. Section IV details the methodological approach employed for this review paper, including the custom web scraper and PRISMA methodology. Section V analyzes the identified research papers. Section VI discusses the chosen research papers, shedding light on the current state, challenges, and prospects. Finally, Section VII concludes the paper by summarizing key findings, contributions, and implications, paving the way for future advancements in transparent and interpretable RSs.

II. RELATED WORKS

The quest for explainability in RSs has spawned numerous research efforts. Given the extensive literature on XRS, a systematic literature review (SLR) mechanism is essential to understand the research paradigm. Multiple surveys and reviews are ongoing to delve deeper into this topic. This section addresses the first research question and delves deeper into the existing surveys and reviews on this topic.

RQ1: What contributions do existing survey or review papers make to the field of XRS when utilizing KG and RL?

This section thoroughly examines eight recent surveys and reviews on the topic, emphasizing the scarcity of reviews focusing on research conducted on XRS using KG and RL. Table 1 summarizes these survey and review papers, primarily identified through the scraper process outlined in the Review Methodology under section IV. Additionally, manual searches on Google Scholar were conducted using the search

TABLE 1. Summary of XRS literature reviews.

Year	Reference	Contribution	Focus Areas	# reviews
2024	[8]	The review primarily concentrates on explainable reinforcement learning (XRL) methods, introducing a novel taxonomy, encompassing a broad spectrum of XRL methods, and offering the necessary granularity to compare various methods and discuss their categories. In addition to reviewing XRL studies, the paper comprehensively summarizes existing XRL surveys and delineates their contributions.	XAI XRL	Reviews (10) Articles (189)
2024	[9]	This survey paper explores advancements and previous integrations of AI and ontologies across diverse fields. The paper proposes ontology-based and semantic-based XAI approaches to enhance the explainability and efficiency of AI outputs through the fusion of machine learning (ML) and semantic reasoning. They particularly emphasize the applicability of these approaches in the manufacturing domain.	XAI KG	Articles (31)
2023	[10]	This study distinguishes between XRL and interpretable ML, emphasizing that XRL goes beyond mere interpretation of decisions. It introduces the Causal XRL Framework, delineating different explanations and evaluating the current research landscape while identifying opportunities for future investigation.	XAI RL	NA
2023	[3]	This survey paper offers a comprehensive overview of deep RL's application in RSs, presenting a classification scheme for existing studies and analyzing them categorically. It also discusses articles that combine KG and RL and provides insights into emerging topics while highlighting promising directions for future research.	XRS RL KG	Articles (66)
2023	[11]	This study delineates a clear distinction between explainability and interpretability within the context of XAI. The paper illuminates the notable challenges and future research directions in XAI. These insights serve as valuable guidance for future explorations in the field of XAI.	XAI RL	Articles (73)
2022	[4]	This paper conducts a systematic survey of KG-based RSs and categorizes the KG-based implementation approaches into three main groups: embedding, connection, and propagation-based methods.	XRS KG RL	Articles (37)
2022	[12]	This review paper categorizes existing literature in the realm of data and knowledge engineering, summarizing their contributions, technical essence, and comparative characteristics. It divides methods into two categories: data-driven, where explanations stem from task-related data, and knowledge-aware, where external knowledge is incorporated.	XAI KG	Articles (42)
2020	[13]	This study offers a comprehensive review of XRS research. It categorizes recommendation problems using the 5W framework (what, when, who, where, and why) and presents a chronological research timeline, from early user studies to contemporary model-based methods. The review summarizes the application of XRS across various tasks, including product, social, and points of interest recommendations.	XAI KG RL	NA

"NA" refers to the non-availability of information.

query enlisting the keywords "survey or review explainable recommender systems by applying knowledge graphs and reinforcement learning." Table 1 provides the publication year of the identified survey or review papers, their contributions, focused areas, and the number of articles and review papers reviewed.

Most existing reviews and surveys focus on XAI, with some exploring implementation variations involving KG and RL. Specifically, Chen et al. [3] presents a survey paper on XRS, primarily focusing on approaches employing deep RL techniques. Guo et al. [4] offers a survey paper on the detailed examination of KG incorporation and a few approaches utilizing RL for XRS development.

Despite the extensive literature on XAI, there remains a noticeable absence of review papers addressing implementation approaches for XRS using KG and RL. Therefore, there is a clear need for a dedicated review paper on this subject. This review paper focuses on providing a comprehensive SLR on XRS using KG and RL. Our approach shares similarities with the work of Pesovski et al. [5], as both studies involve reviewing the literature on XRS. However, instead of using an NLP-powered toolkit, we opted for a Selenium and Python-based web scraper to programmatically extract articles from major online digital libraries, including IEEE, ScienceDirect, ACM, and Springer. Our review aims to comprehend the broader spectrum of XRS.

III. BACKGROUND

This section provides conceptual knowledge before diving deep into the review methodology.

A. RECOMMENDER SYSTEMS

RSs are algorithmic frameworks designed to predict and suggest items or content that users may perceive as pertinent or appealing. These systems leverage various data sources, including user preferences, historical interactions, and item attributes, to generate tailored recommendations [6]. With widespread application across diverse domains such as e-commerce, streaming services, and social media, RSs aim to enrich user experiences, boost engagement, and drive desired outcomes like sales or user interactions [7]. As digital platforms continue to evolve, RSs play a pivotal role in shaping user experiences by dynamically predicting and delivering personalized content recommendations. From their origins in content-based and collaborative filtering-based models, RSs continuously adapt to accommodate the multi-faceted preferences of users [14], [15].

B. EXPLAINABILITY IN RECOMMENDER SYSTEMS

RQ2: What defines explainability in artificial intelligence (AI), and why is it important for RS?

RSs are extremely helpful and assist users in their tasks. However, concerns arise about user trust due to the prevalent use of opaque algorithms, prompting a demand for transparency. In response, XRS has emerged, focusing on accurate recommendations and providing understandable justifications. This justification addresses users' need for comprehension and trust in algorithmic decision-making processes. This context forms the basis for thoroughly exploring XRS intricacies, where transparency and interpretability intertwine with RS dynamics [1]. The subsections below provide an overview of explainability.

1) WHAT IS EXPLAINABILITY?

Explainability involves clarifying the functionality of a model and its predictions, recommendations, and suggestions to users and decision-makers. It aims to answer questions like "How does the AI system work?" and "Why are specific recommendations made?" It illuminates the rationale behind recommendations and helps with suitable explanations. It also helps identify biases and weaknesses in the system, boosting user confidence and trust [16].

Fig. 2 illustrates the XAI process, which begins with users receiving initial predictions or recommendations from the AI model. At this stage, users form an initial mental model of both the task and the AI system, which may include mistrust towards the model's predictions. However, subsequent experiences, including explanations provided by the system, allow users to refine their mental model over time, potentially leading to improved performance and a greater sense of trust and reliance on the AI system [16].

In this process, explainability is crucial in fostering transparency and trust among users. Effective and satisfactory explanations assist users in developing a more comprehensive and accurate mental model of the AI system and its predictions. This well-formed mental model leads to increased trust in the AI system. As trust strengthens, users will interact with the system more adeptly. Consequently, better explainability enhances transparency and contributes to users' evolving understanding, enabling them to make more informed decisions and effectively utilize the capabilities of the AI system [16].

2) THE NEED FOR EXPLAINABILITY

The significance of explainability in AI systems stems from various factors, including ethics, regulatory compliance, and transparency, to infuse trust in the minds of the end-users and business stakeholders. This emphasis arises from concerns about transparency, accountability, bias mitigation, and ethical considerations.

AI ethics advocates for fairness and accountability in AI systems, ensuring stakeholders can identify biases, errors, or unethical behaviour. This accountability promotes responsible AI development [17]. Regulations like GDPR [18], the European Union AI Act 2023 [19], and the California Consumer Privacy Act of 2018 [20] emphasize the "right to explanation" in algorithmic decision-making, highlighting the importance of AI system explainability. Providing explanations enhances user trust by offering insight into decision-making processes and fostering confidence in critical domains like healthcare or finance [21]. The combination of these reasons underscores the need for transparent, ethical, and regulated AI systems, marking the advent of responsible AI.

C. KNOWLEDGE GRAPH

While the concept of an intelligent model encoding real-world entities and their relationships has been present in literature since the 1980s [22], the term KG gained popularity following Google's announcement in 2012. Since then, various KGs have emerged, including Freebase [23], DBpedia [24], YAGO [25], and Google's own KG [26], facilitating the construction of KGs for RS. Despite its widespread use, a precise definition of KG remains elusive [27]. However, there is a general understanding of its key characteristics.

Formally, a KG Gknow $= (V, E)$ is a directed graph with V vertices (nodes) and E edges (relationships). Nodes represent entities, while edges denote subject-predicate-object triplet facts. Each edge of the form $<$ e_h $(headentity)$ $,r(relation),e_t(tailentity$ $>$ indicates a relationship r, from the head entity e_h to the tail entity e_t [4]. A KG serves as a structured representation of information, portraying relationships between real-world entities through a directed graph. Nodes symbolize entities, and edges depict relationships between these entities, often labelled to offer context or additional information [22]. Furthermore, the KG's structure may include an ontological schema, organizing entities and relationships hierarchically or categorically to provide a more structured view of the information within the graph [28].

A KG has two fundamental components [29]:

- Terminology Box (T-Box) – A KG includes a set of concepts, categories, and properties that establish the foundational framework for understanding the entities and relationships within the graph.
- Assertion Box (A-Box) – A KG includes a set of statements or assertions about individual entities, specifying their attributes, relationships, and other relevant information. These statements, known as the A-Box, contain specific data instances that align with the concepts and relationships defined in the T-Box.

A KG aims to provide a semantically rich and interconnected representation of data. It is used to enhance search engines,

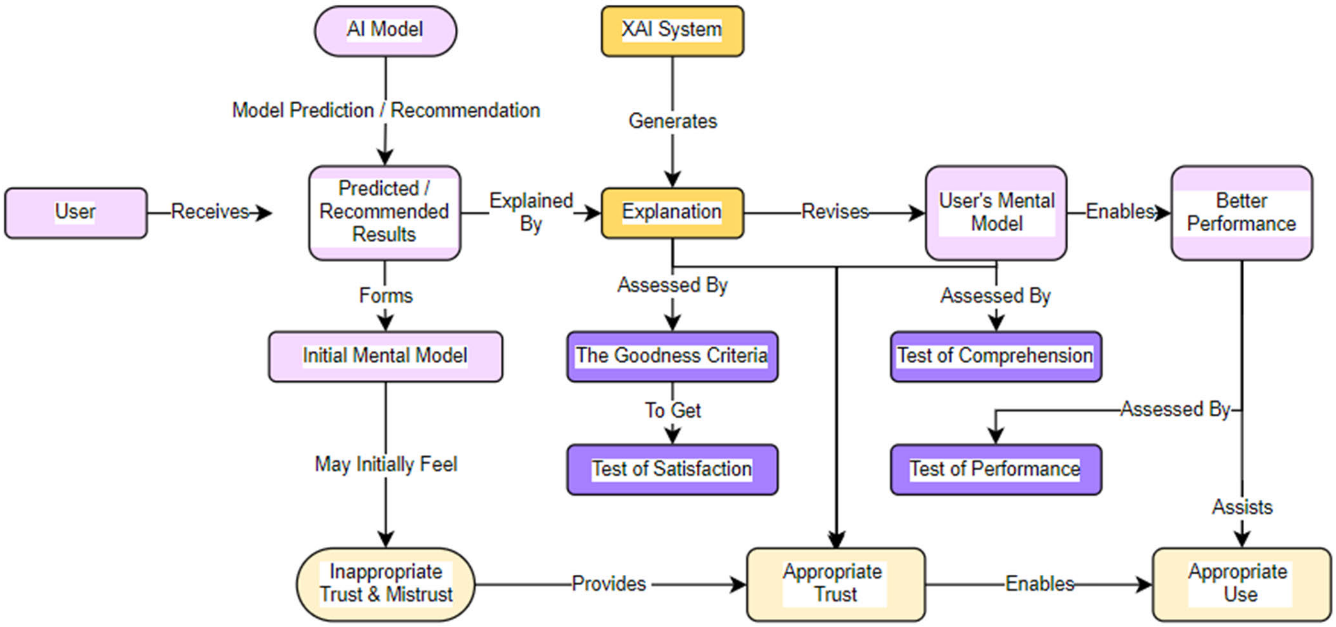

FIGURE 2. A conceptual model of explaining in the XAI context.

facilitate machine reasoning, support question-answering systems, and more by enabling a deeper understanding of the relationships between various entities and concepts in the real world.

D. REINFORCEMENT LEARNING

The RL framework comprises core elements: Environment-State, Agent, Actions, and Rewards. An autonomous agent improves its performance by interacting with its environment. As described by [30], an agent with sensors and actuators evaluates its performance based on a reward function. It navigates through states, selecting actions and receiving rewards periodically. The agent's primary goal is to learn a policy that maximizes cumulative rewards over time and balances exploration and exploitation. It exploits current knowledge while exploring new possibilities [31]. Key components include states, actions, rewards, value functions, and Q-functions, all integrated within the Markov Decision Process (MDP) framework. Primary RL algorithms include Q-learning, REINFORCE, and Actor-Critic.

1) MARKOV DECISION PROCESS

MDP is considered the de facto standard when formalizing sequential decision-making problems involving an RL agent, in which actions affect the current short-term rewards, the subsequent states, and future rewards [32]. An MDP represents a tuple $< S, A, T, Rw, \gamma >$, where S is the set of all possible states, which is the generalization of the environment, A is the set of all possible actions of the agent that can be adopted in the states, T is a transition function, Rw is a reward function, and γ is a discount factor [33].

2) Q-LEARNING

Q-learning [34] is an off-policy and Temporal Difference (TD) learning algorithm. It operates by updating values within a Q-table using a target policy to attain the optimal policy. Meanwhile, a separate behaviour policy, often employing the ε-greedy strategy, facilitates semi-random exploration of the environment. The primary objective of Q-learning is to approximate the optimal action-value function, $q*$, through direct estimation, representing the highest expected cumulative reward achievable for actions in specific states.

3) REINFORCE

REINFORCE [35] is an on-policy algorithm that directly learns a parameterized policy using the Monte Carlo method. It updates policy parameters based on estimated returns and full traces. This method employs neural networks to construct a policy, with states as inputs and outputting probability distributions in the operational space.

4) ACTOR-CRITIC

The Actor-Critic algorithm [36] in RL combines a parameterized policy (Actor) and a value function (Critic) to improve learning efficiency and stability. The Critic estimates the expected cumulative reward for a state-action pair, guiding policy optimization through policy gradient methods, enabling the Actor to make more informed decisions. By integrating policy-based and value-based techniques, the algorithm balances exploration for maximizing rewards with an accurate assessment of state-action desirability.

This review paper conducts an SLR encompassing all related articles applying KG and RL. Sections IV, V, and VI offer a comprehensive overview of the review process and discussions, comprehend the research paradigm surrounding KG and RL in the context of XRS, and address the research questions RQ3, RQ4, and RQ5.

IV. REVIEW METHODOLOGY

This section outlines the review methodology, covering the use of a web scraper and the PRISMA methodology.

A. WEB SCRAPER – SCRAPE THE ARTICLES

In the vast realm of academic research, obtaining research papers from reputable online databases poses a considerable challenge and requires manual human effort. This study introduces a custom web scraper, as depicted in Fig. 3, to simplify accessing and organizing research papers from respected online databases like ACM, Springer, IEEE Xplore, and ScienceDirect. Built with Python, the scraper uses the Microsoft EDGE web driver and Selenium for automated web scraping. It helps researchers efficiently find and manage academic content. The custom web scraper framework is available on the GitHub link shared on the first page.

The efficacy of the web scraper is contingent upon the precision of the search query, which acts as the primary filter for article retrieval and inclusion. It systematically evaluates article titles, abstracts, and keywords for specific search terms, ensuring comprehensive data retrieval across multiple pages. By adapting its operations based on the number of pages, the scraper meticulously reviews the results of each page and selectively stores those that meet the specified criteria.

The extracted results contain essential information, including the article's web link (URL), title, author, publication year, journal or conference name, online database source, citation count, accessibility status, original search query, and other search criteria. This comprehensive set of results offers a detailed overview of the identified articles, making it easier to analyse and explore these relevant publications.

Table 2 presents the queries used in this study to identify relevant articles. These queries yielded 130 articles from ACM, 257 articles from Springer, seven research papers from IEEE, and five articles from ScienceDirect. This comprehensive search process returned 399 articles, providing the foundation for further refinement.

B. PRISMA – PRIORITISE THE SCRAPED ARTICLES

PRISMA methodology, or "Preferred Reporting Items for Systematic Reviews and Meta-Analyses", serves as a standardized framework for systematic reviews and surveys, improving practices for researchers conducting literature reviews and enabling the unbiased study of published articles [2]. The process begins with identifying articles based on specific criteria, such as keywords, followed by screening, eligibility assessment, and inclusion of articles aligned with the survey paper's objectives. This systematic approach ensures the thorough selection and processing of relevant articles for this review paper on XRS utilizing KG and RL.

The methodology illustrated in Fig. 4 was employed to refine the article selection process. Initially, our search criteria yielded 399 articles from reputable databases such as ACM, Springer, IEEE Xplore, and ScienceDirect. Following a thorough screening, 354 articles, accounting for 88.7% of the total, were retained. Subsequently, 10.7% of the articles met the eligibility criteria, resulting in 38 articles. Further refinement during the final inclusion stage led to 29 articles aligned with the specified search criteria and filtering options.

1) STEP 1: IDENTIFICATION OF ARTICLES

The study examined research papers across major online databases, including ACM, Springer, IEEE Xplore, and ScienceDirect, focusing on the explainability of RSs utilizing KG and RL. The search query "Explainable AND (Recommendation OR Recommender) AND Knowledge Graph AND Reinforcement Learning" was employed to identify relevant research papers from January 2015 to March 2024. Only research papers written in English, with the query keywords in the document title, abstract, and keywords, were included.

2) STEP 2: SCREENING OF IDENTIFIED ARTICLES

The next step involves refining the candidate articles from the previous identification phase. Since the articles are collected from various online database sources, there is a chance of duplicate entries or irrelevant articles unrelated to the review objective. The articles undergo manual screening to identify duplicity or irrelevance to the core subject, such as articles with prefaces or editorial content, resulting in the identification of 354 articles.

3) STEP 3: ELIGIBILITY OF SCREENED ARTICLES

This step involves identifying eligible articles from the previous screening phase. We review all the previously screened articles to determine their purpose and underlying technologies. We refine the list of potential papers by identifying the articles focusing on developing XRS using KG and RL and exclude those confined to review or survey purposes. This process yields 38 core articles.

4) STEP 4: INCLUSION OF ELIGIBLE ARTICLES

In the concluding step, we further refine the candidate papers to align them precisely with the core objective of the review paper. We exclude articles from the eligible list of research that predominantly address temporal research and counterfactual studies or are unrelated to the core development of XRS using both KG and RL. Additionally, we include any other relevant research that may have been overlooked during the scraping process. This last step ensures a focused set of 29 research papers for the further review process.

V. RESULTS ANALYSIS

During this phase, a thorough analysis of the scraped articles obtained in the previous stage is conducted, aiming to observe patterns in research related to explainable recommenders across various platforms within the research community. This

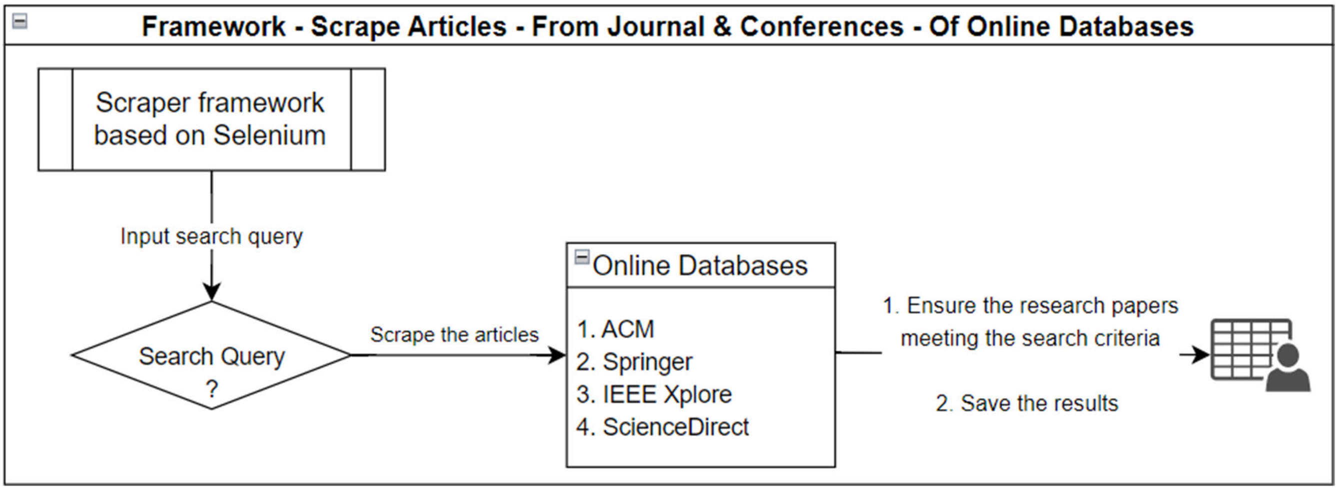

FIGURE 3. Web scraper framework.

TABLE 2. Search queries used for different online databases.

Database	Query Used
ACM	https://dl.acm.org/action/doSearch?fillQuickSearch=false&target=advanced&expand=dl&AfterYear=2015&BeforeYear=2024&AllField=Title%3A%28Explainable AND (Recommendation OR Recommender) AND Knowledge Graph AND Reinforcement Learning%29+OR+Abstract%3A%28Explainable AND (Recommendation OR Recommender) AND Knowledge Graph AND Reinforcement Learning%29+OR+Keyword%3A%28Explainable AND (Recommendation OR Recommender) AND Knowledge Graph AND Reinforcement Learning%29&startPage=&ContentItemType=research-article
Springer	https://link.springer.com/search?new-search=true&query=Explainable AND (Recommendation OR Recommender) AND Knowledge Graph AND Reinforcement Learning&content-type=Article&content-type=ConferencePaper&date=custom&dateFrom=2015&dateTo=2024&language=En&sortBy=relevance
IEEE	https://ieeexplore.ieee.org/search/searchresult.jsp?action=search&matchBoolean=true&queryText=(%22Document%20Title%22:Explainable AND (Recommendation OR Recommender) AND Knowledge Graph AND Reinforcement Learning)%20OR%20(%22Abstract%22:Explainable AND (Recommendation OR Recommender) AND Knowledge Graph AND Reinforcement Learning)%20OR%20(%22Author%20Keywords%22:Explainable AND (Recommendation OR Recommender) AND Knowledge Graph AND Reinforcement Learning)&highlight=true&matchPubs=true&returnType=SEARCH&ranges=2015_2024_Year&returnFacets=ALL&refinements=ContentType:Conferences&refinements=ContentType:Journals&refinements=ContentType:Early%20Access%20Articles
Science Direct	https://www.sciencedirect.com/search?date=2015-2024&tak=Explainable AND (Recommendation OR Recommender) AND Knowledge Graph AND Reinforcement Learning&langs=en&articleTypes=FLA%2CREV

analysis contributes to a deeper understanding of trends and insights within the realm of explainable recommenders across diverse research platforms.

A. QUANTITATIVE ANALYSIS

This section conducts the quantitative analysis of the scraped articles.

1) ANALYSIS OF THE PUBLISHED SCREENED ARTICLES BY YEAR

Fig. 5 illustrates the proliferation of published articles retrieved using the search query used in the previous section across all four reputable online databases from January 2015 to March 2024. These articles underwent screening to eliminate irrelevant ones. Notably, the search query only returned articles from 2017 onwards, as no relevant articles were found for 2015 and 2016 based on our search criteria. Over the years, the rising number of published articles underscores the increasing interest within the research community in XRS utilizing KG and RL. In 2023, for instance, 125 articles were published, showing the community's interest in XRS and related topics.

2) ANALYSIS OF PUBLISHED ELIGIBLE ARTICLES BY ONLINE DATABASES

Fig. 6 depicts the distribution of eligible articles across different online databases, highlighting the results of a manual critical review process to assess their relevance to the research objectives. Springer and ACM stand out as the primary online databases, containing a substantial number of

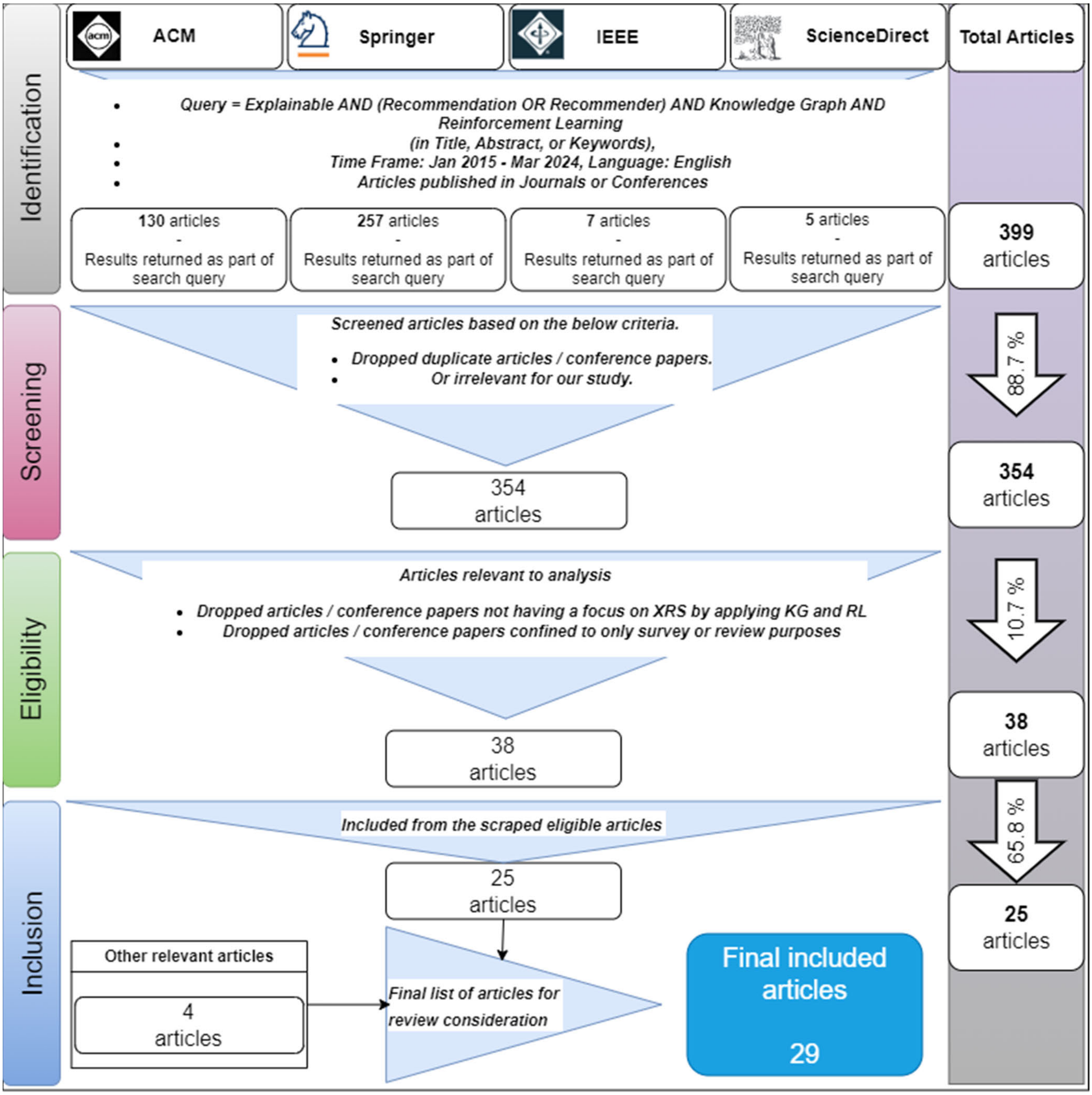

FIGURE 4. The PRISMA methodology of identifying the research papers for reviewing.

published articles relevant to the review objectives. Despite initially identifying only a few articles from IEEE and ScienceDirect, these sources remained relevant throughout the eligibility process.

3) ANALYSIS OF PUBLISHED SCREENED ARTICLES BY YEAR BY OPEN-ACCESS STATUS

Fig. 7 illustrates the prevalence of open-access status among the screened articles, showing a gradual increase since 2018 and a significant surge from 2020 onwards. There was a notable increase in 2023, with 51 articles published as open access compared to 76 as non-open access. The years 2022, 2023, and 2024 witnessed a considerable rise in the proportion of open-access articles, exceeding 40%. As of 2024, open-access articles represent 43.6% of the total. This trend underscores the importance of freely sharing research findings to enhance accessibility and collaboration in the research community.

B. QUALITATIVE ANALYSIS

This section conducts the qualitative analysis of the articles included in the review process.

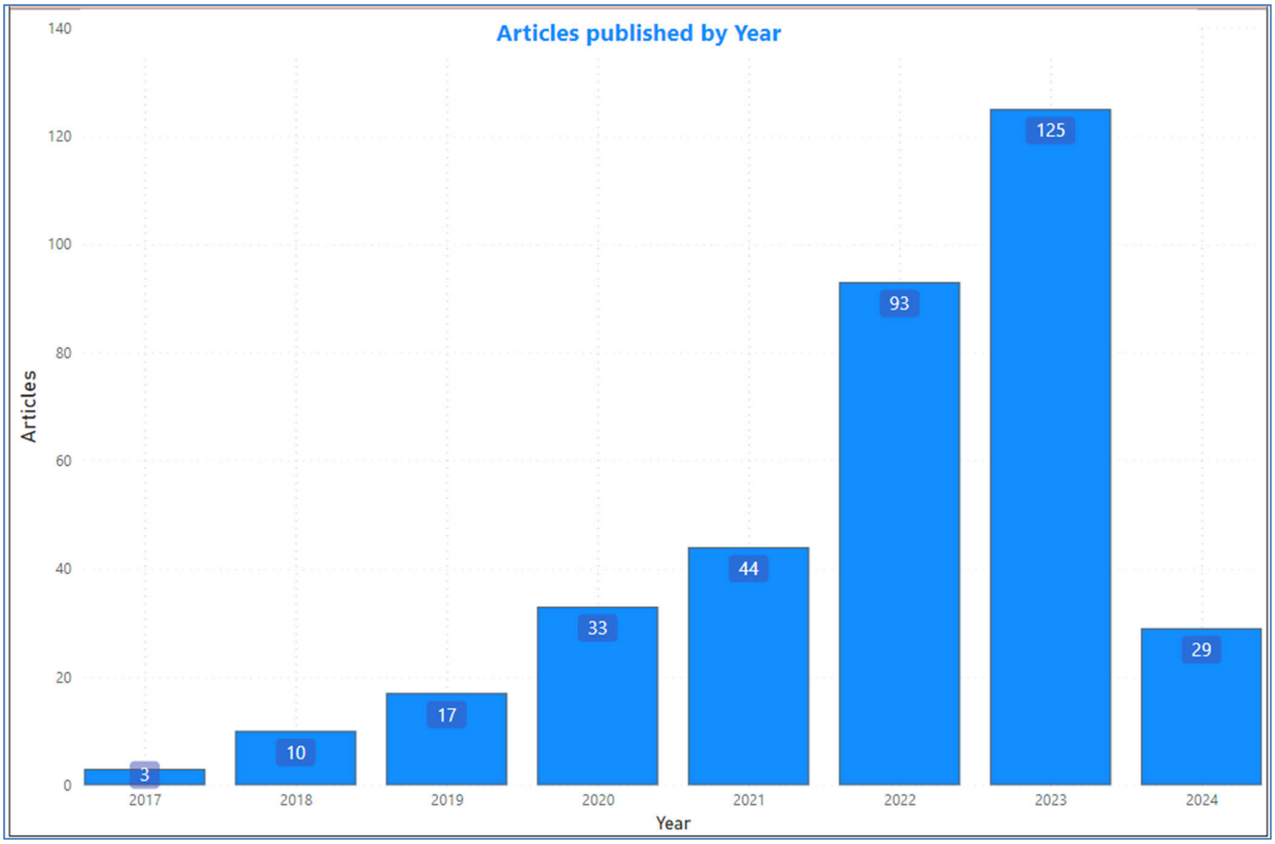

FIGURE 5. Screened articles published by year.

This sub-section highlights 29 seminal research works in XRS utilizing KG and RL. Table 3 provides a comprehensive overview of these selected studies, detailing their publishing platforms, explainability methodology, applied technologies, experimental domains, datasets, and approaches. The selection process of these 29 articles involves the scraper framework and the PRISMA methodology, resulting in the 25 pertinent articles, and the manual search mechanisms contribute to including the remaining four articles.

The final selection of review articles covers XRS using KGs and RL across diverse domains, spanning major online databases like IEEE, Springer, ACM, and ScienceDirect. The distribution includes 10 ACM articles, eight from Springer, eight from IEEE, and three from ScienceDirect. KGs and RL are pivotal technologies enabling model-intrinsic or embedding explainability for recommendations. As outlined in the discussion section, the approaches vary in algorithmic adoption and implementation.

VI. DISCUSSIONS

This section delves into a detailed discussion of the included research works and relevant methods for XRS development and addresses research questions RQ3, RQ4, and RQ5.

A. METHODS OF EXPLAINABLE RECOMMENDER

The rising interest in XRS stems from its potential to enhance user acceptance of recommendations by providing transparent explanations. Fig. 8 outlines various approaches to implementing XRS, aiming for XAI. These approaches are categorized into model-agnostic or post-hoc and model-intrinsic or embedded methods. KG-based and RL-based approaches are the primary strategies. KG-based methods include embedding-based, path-based, and hybrid, while RL-based methods comprise value-based, policy-based, and hybrid approaches. While explainability can be achieved without KG and RL, those methods are mostly model agnostic. Integrating KG and RL techniques presents new methodological advancements. These integrations leverage structured knowledge representation (KG) and learning through interaction (RL) to enhance the transparency and effectiveness of RSs. This section conducts a comprehensive analysis of the scraped articles.

B. EXPLAINABILITY METHODOLOGY

The explainability of the XAI can be categorized into model-agnostic and model-intrinsic.

1) MODEL-AGNOSTIC OR POST-HOC METHODOLOGY

Post-hoc methodology involves generating explanations after the recommendation process, often relying on previously identified sets of candidates, such as "people also bought" or "7 of your friends like this." While these explanations are typically persuasive and easy to understand, they overlook the inner workings of the RS.

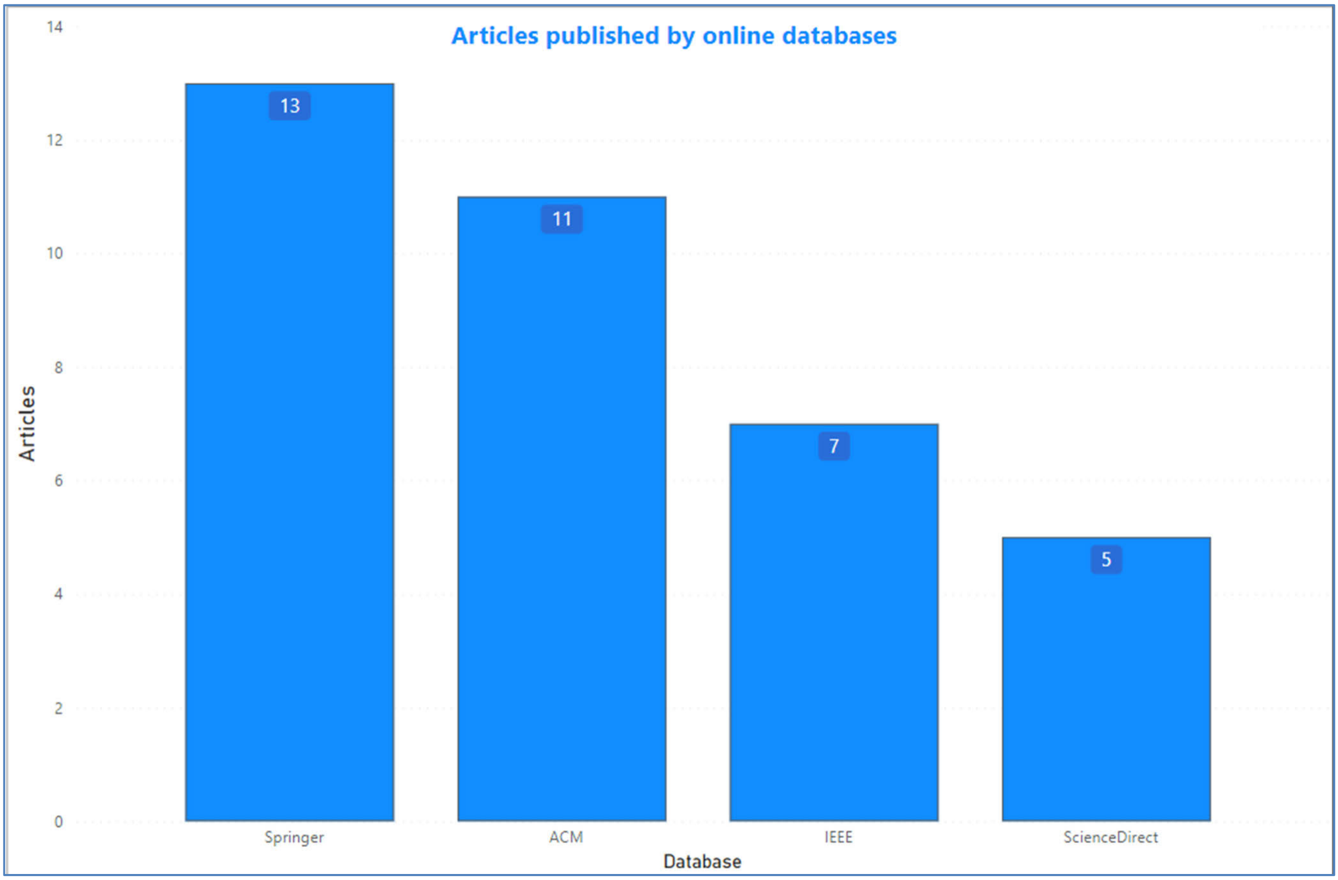

FIGURE 6. **Eligible articles published by online databases.**

Ai et al. [66] employ a CF approach over KGE, using a soft matching algorithm to find explanation paths between users and items, resulting in a post-hoc explanation method. However, this approach overlooks the inner workings of the recommendation model. Valdiviezo-Diaz et al. [67] demonstrate the effectiveness of leveraging the Naïve Bayes Classifier to enhance RS design, improve recommendation quality, and provide transparent justifications. Their NBCF approach involves developing a probabilistic model using the Naïve Bayes Classifier and adopts a post-hoc methodology for explainability. Jung and Nardelli [68] introduce a model-agnostic explainability model by explicitly modelling user-specific knowledge. They propose a probabilistic model for data and user background, measuring the quantitative effects of explanations through mutual information. Unlike other model-agnostic approaches, this method employs a probabilistic model for predictions and user knowledge, considering personalized explanations.

While post-hoc methods are model agnostic, they are not preferable because they do not consider model explainability when generating diversified results.

2) MODEL-INTRINSIC OR EMBEDDED METHODOLOGY

Embedded methods integrate the explanation process into RS construction, deriving explanations from item-side information to emphasize factors contributing to recommendation accuracy. While these methods produce diverse, model-tied explanations, ensuring readability and consistency may be challenging. KG-based and RL-based solutions are prominent implementation approaches, detailed in the following sections.

C. KG-BASED XRS

This section provides an overview of various methodologies focusing on KG-based approaches for XR, tackling research question RQ3. These methods are classified into embedding, connectivity, and hybrid.

1) EMBEDDING-BASED METHODS

Embedding-based methods typically leverage KG information to enhance item or user representations. KGE algorithms encode KGs into low-rank embeddings. These algorithms fall into two categories: translation distance models (e.g., TransE [69], TransH [70], TransR [71], TransD [72]) and semantic matching models (e.g., DistMult [73]). KGE models determine entity similarity by computing their representation distance [74].

Several embedding-based methods [75] construct KGs with multiple item side information to enrich item representations, forming item graph-based KGs that improve user

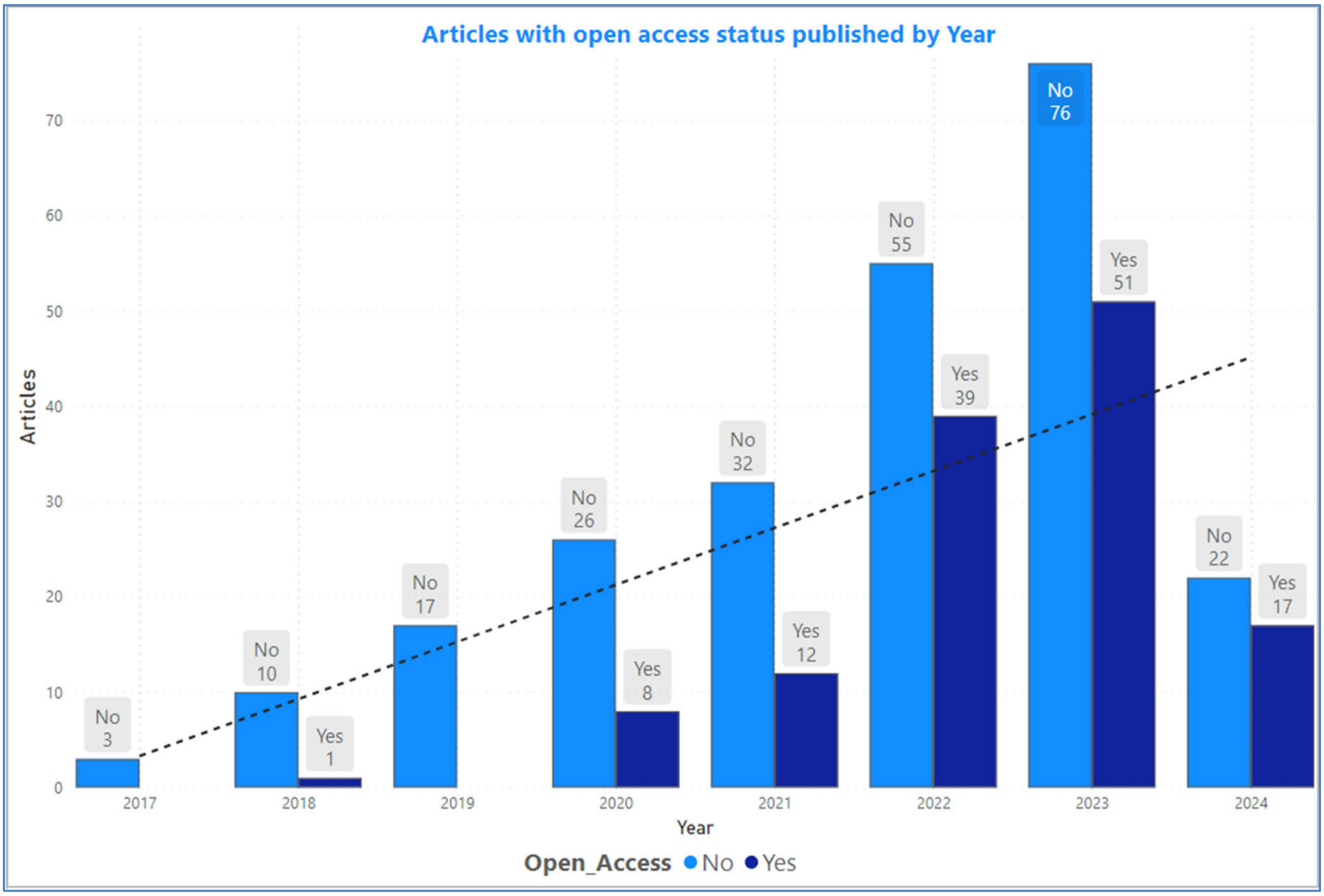

FIGURE 7. Screened articles published by year by open-access status.

modelling. Some models [66] incorporate user preferences through user-item KGs by integrating users into the graph design. Entity embedding is fundamental to these methods, with few studies refining embeddings using techniques like GAN [76] or BEM [77] to improve recommendations. The problem with these methods is overlooking multi-hop relational paths.

2) CONNECTIVITY OR PATH-BASED METHODS

Path-based recommendation within KG methodologies leverages user-item graphs to enhance recommendations by exploiting connectivity patterns, thereby improving recommendation quality.

Prominent research works on path-based methods include Hete-MF [78], Hete-CF [79], HeteRec-p [80], MCRec [81], and RKGE [82], which utilize meta-paths and their variations for different users. Zhao et al. [83] introduced FMG, replacing the meta-path with the meta-graph. Additionally, Wang et al. [84] proposed KPRN, which constructs the extracted path sequence with entity and relation embeddings, encoding it with an LSTM layer.

Path-based methods utilize meta-paths to match item or user similarity, enhancing interpretability. Recent advancements employ deep learning models to extract significant traversal paths for user-item pairs.

3) HYBRID OR UNIFIED METHODS

Embedding-based methods utilize the KG's semantic representation of users/items, while path-based methods rely on semantic connectivity information. However, each approach only utilizes one aspect of KG information. Unified methods integrate both semantic entity representation and connectivity information, leveraging embedding propagation. This approach refines entity representation with KG's connective structure guidance.

The prominent unified-based methods, including RippleNet [85], AKUPM [86], RCoLM [87], KGCN [88], KGAT [89], IntentGC [90], and AKGE [91], leverage preference propagation. Ryotaro [92] introduces an enhanced KGAT model, reducing computational costs by compressing extensive side information while preserving high accuracy and explainability. Thus, the unified methods blend KG semantic embedding and path patterns, refining item or user representation through multi-hop embedding propagation. They inherit interpretability from path-based methods, revealing user preference patterns akin to connectivity patterns.

TABLE 3. Comparison of xrs research works utilizing KG and RL.

S. No	Model	Publish Site	Explainability	Techniques	Domain/ Dataset	Approach
1	UPGPR [37]	ACM	Model-intrinsic	KG RL - Policy	Educational/ COCO and Xuetang	The study proposes an XRS for MOOCs that uses RL applied on KG reasoning. They also conducted a user study examining user perceptions of their XRS.
2	RCENR [38]	ACM	Model-intrinsic	KG RL – Hybrid	News/ MSN-MIND	The study proposes an explainable news recommendation model consisting of user/news KG, contrastive, and RL to solve data sparsity and poor interpretability problems.
3	DPAO [39]	ACM	Model-intrinsic	KG GNN RL - Value	E-commerce/ Yelp2018, Last-FM, etc.	The study proposes an adaptive aggregation strategy for user- and item-side Graph Neural Networks (GNNs) and introduces an RL-based Dual Policy framework for Aggregation Optimization (DPAO) using Deep-Q-Network (DQN) models.
4	ReMR [40]	ACM	Model-intrinsic	KG RL - Hybrid	Amazon-Beauty, Cloth, etc.	The study proposes an RL framework for a multi-level recommendation over KGs and leverages ontology-view and instance-view KGs to model multi-level user interests.
5	RMS [41]	ACM	Model-intrinsic	KG RL - Value	Yelp, Douban, TCL	The study proposes an RL-based meta-path selection framework RMS to find an effective meta-path set for automatic recommendation.
6	Anchor-KG [42]	ACM	Model-intrinsic	KG RL - Hybrid	News/ MSN-MIND, Bing-News	This study proposes to develop a compact anchor KG for each article and leverages an RL-based framework to train the anchor graph generator.
7	ADAC [43]	ACM	Model-intrinsic	KG RL - Hybrid	E-commerce/ Beauty, Cell Phones, Cloth	The study proposes a demonstration-based KG reasoning framework for XR and introduces an ADversarial Actor-Critic (ADAC) model for demonstration-guided pathfinding.
8	PGPR [44]	ACM	Model-intrinsic	KG RL - Policy	E-commerce/ Beauty, Cell Phones, CDs, Cloth	This study proposes Policy-Guided Path Reasoning (PGPR) to enhance recommendation and interpretability by utilizing an RL approach with a soft reward strategy, user-conditional action pruning, and a multi-hop scoring function.
9	SAPL [45]	ACM	Model-intrinsic	KG RL - Policy	Amazon – Cellphones, Cloth, & CD	This study introduces SAPL, a Sentiment-Aware Policy Learning, integrating sentiment analysis with KGs for accuracy and explanation. It proposes a Sentiment-Aware KG (SAKG) based on user reviews and ratings.
10	KGQR [46]	ACM	Model-intrinsic	KG RL - Value	Book-crossing, ML20M	This study introduces a method leveraging prior item correlation knowledge from KG to guide candidate selection for improved item retrieval, enhances item and user state representations and propagates user preferences to address feedback sparsity.
11	CERec [47]	IEEE	Model-intrinsic	KG RL - Policy	Last-FM, Yelp2018, Amazon-Book	This study proposes a Counterfactual XR (CERec) to generate item attribute-based counterfactual explanations within an RL environment to boost recommendation performance.
12	REKS [48]	IEEE	Model-intrinsic	KG RL - Policy	Amazon-Cell Phones, Baby Beauty, etc.	This study presents REKS, a Reinforced Explainable framework with a KG for Session-based recommendation (SR), enhancing accuracy and explainability using RL.
13	TMER-RL [49]	IEEE	Model-intrinsic	KG RL - Policy	Amazon - Music Inst. Automotive	This study introduces TMER-RL, a Temporal Meta-path Guided XR leveraging RL and leveraging supervised RL to explore item-item paths in a dynamic KG for an XR.
14	RKGR-RNS [50]	IEEE	Model-intrinsic	KG RL - Policy	Amazon-Cell Phones, Beauty, Cloth	This study introduces RKGR-RNS, a Reinforced KG Reasoning with a Reinforced Negative Sampling approach, incorporating a negative sampling method into RL-based recommendation.
15	PPO [51]	IEEE	Model-intrinsic	KG RL - Policy	Amazon-Cell Phones, Beauty	This study proposes to use Proximal Policy Optimization (PPO), an RL method, to optimize the path exploration problem.
16	Mcore [52]	IEEE	Model-intrinsic	KG RL - Policy	Amazon- CD, Cell Phones, Beauty, Cloth	This study introduces Mcore, a novel Multi-agent Collaborative RL framework designed to enhance model performance by extracting users' high-level interests. It is adaptable to large KGs.

TABLE 3. ***(Continued.)*** **Comparison of xrs research works utilizing KG and RL.**

17	KGRL [53]	IEEE	Model-intrinsic	KG RL - Hybrid	Amazon CD, LibraryThing, ML20M, Netflix, etc.	This study introduces KGRL, which integrates RL and KG into an actor-critic network for interactive recommendation. It utilizes a local knowledge network for decision-making guidance and attention mechanisms to capture long-term item semantics.
18	KGDQN [54]	IEEE	Model-intrinsic	KG RL - Value	Amazon-Cellphones, & Beauty	This study combines KG and RL for recommendations, pruning redundant edges and using a DQN model to provide recommended items and explanation paths.
19	MES [55]	Science Direct	Model-intrinsic	KG RL - Policy	E-commerce/ Amazon-CD, Cell, Music, Book, Beauty	This study proposes MES, Max Explainability Score, a novel metric for quantitative evaluation of explanation quality in KG-based RSs based on the number of rules, the probability, the entropy value, and the reward of chosen traversal paths.
20	IP-PGPR [56]	Science Direct	Model-intrinsic	KG RL - Policy	MovieLens1M (ML1M), LastFM-1B, A-Cellphones	The study introduces quantitative properties for monitoring reasoning path quality, integrating recency, popularity, and diversity metrics. It employs both in- and post-processing techniques to enhance recommendation and reasoning quality.
21	TN [57]	Science Direct	Model-intrinsic	KG RL - Policy	NELL-995, Amazon Beauty & Cellphones	The study presents a type-enhanced RL agent leveraging local neighbourhood information for efficient path-based reasoning on KGs. It employs a GNN for encoding neighbourhood information and utilizes entity types for action space pruning.
22	RSL-GRU [58]	Springer	Model-intrinsic	KG RL - Policy	E-commerce/ Amazon-Movie, Cloth, Beauty, Book	This study proposes RSL-GRU, a Reinforced Sequential Learning with a Gated Recurrent Unit composed of a Reinforced Path Reasoning Network (RPRN) component and a gated recurrent unit (GRU) component.
23	KRRL [59]	Springer	Model-intrinsic	KG RL - Hybrid	Education/ MOOC-Course, MOOC-Cube	The study presents KRRL, a method using self-supervised RL for Knowledge-aware Reasoning to improve semantic representation and relationships in the KG, employing multi-level representation learning for enriched semantic interactions.
24	KPRLN [60]	Springer	Model-intrinsic	KG RL - Value	ML1M Last.FM	The study presents KPRLN, a deep knowledge preference-aware RL network. It establishes paths between user historical interactions in the KG, learns preference features, and generates a weighted KG with fine-grained preferences.
25	IMRL [61]	Springer	Model-intrinsic	KG RL - Value	Netflix Prize, MoveLens-20M, Book-crossing, etc.	The study introduces IMRL, an intrinsically motivated RL method to enhance the agent's exploration in sparse environments through an adaptive counterfactual strategy with a customized threshold for improved efficiency in exploitation.
26	HINRL4 Rec [62]	Springer	Model-intrinsic	KG RL - Policy	ML-100 K Amazon-Baby, Music, Book, Beauty	The study proposes HINRL4Rec, integrating HIN embedding with RL-based feature engineering for recommendation using textual meta-path-based embedding fused into unified spaces, facilitating RL-based policy-driven searching for XR.
27	MEUR [63]	Springer	Model-intrinsic	KG RL - Hybrid	Education/ MOOCCube	The study proposes MEUR, Multi-path Embedding and User-centric Reasoning. This resource recommendation method embeds multiple paths and searches with users as the centre, innovatively combining the advantages of GCN and RL.
28	CDARL [64]	Springer	Model-intrinsic	KG RL - Policy	E-commerce/ RC15 RetailRocket	The study introduces CDARL, a contrastive discriminator-augmented RL framework with a contrastive learning module for item sequence representation and an intrinsic reward learning function to mimic user internal dynamics.
29	HR-RL-KG [65]	Springer	Model-intrinsic	KG RL - Policy	E-commerce/ Amazon-CD, Cell, Music, Book, Beauty	This study introduces HR-RL-KG, which integrates RL and KG. It introduces a novel reward function for flexible reward calculation and employs the TransD method for vector embedding of the KG.

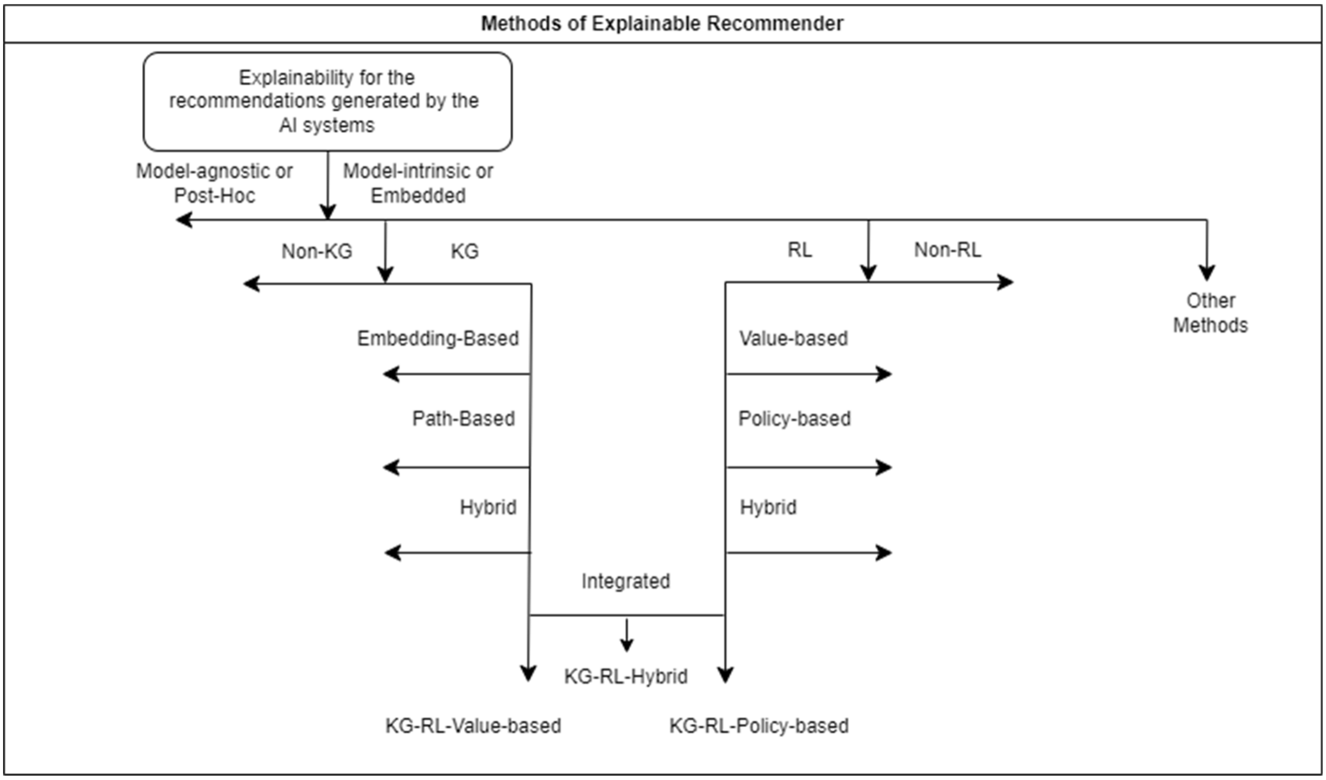

FIGURE 8. Methods of explainable recommender.

TABLE 4. KG-based XRS research works.

KG-based approach	Key XRS research works mentioned in this review paper
KGE-based methods	[69], [70], [71], [72], [73], [74], [75], [66], [76], [77]
Path-based methods	[78], [79], [80], [81], [82], [83], [84]
Hybrid methods	[85], [86], [87], [88], [89], [90], [91], [92]

Table 4 summarizes the KG-based research works mentioned in this review paper. This subsection discusses various KG-based implementation approaches comprising embedding-based, path-based, and hybrid methods.

Embedding-based methods harness the semantic representations of users/items in the KG, while path-based methods utilize semantic connectivity data. Unified methods combine both aspects, fully exploiting KG information by integrating entity and relation semantic representations with connectivity data. They leverage embedding propagation, refining entity representations based on KG connective structures.

RQ3: How does KG contribute to achieving explainability within RS?

KGs contain relationships between entities, enabling the utilization of related information to explain recommendations. They comprehend user interests by understanding entity semantics and relations, providing rich data beyond user-item interactions. KGs alleviate cold-start issues in RS. Recommendation reasoning is categorized as Ante-hoc or Post-hoc. Ante-hoc involves model-intrinsic explainability, while Post-hoc solutions are model agnostic, providing explanations after recommendations. While KGs provide recommendation reasoning, several challenges remain in developing robust solutions that accurately capture user preferences.

D. RL-BASED XRS

This section provides an overview of various methodologies focusing on RL-based approaches for XR to address research question RQ4. These methods are classified into value-based, policy-based, and hybrid.

1) VALUE-BASED METHODS

Value-based RL involves the agent learning to make decisions by estimating the value of different actions in various states. This estimation creates a value function, assigning values to states or state-action pairs. The agent's objective is to maximize cumulative rewards over time by choosing actions leading to states with higher estimated values. Algorithms like Q-learning and DQN update these estimates iteratively through interaction with the environment, adjusting based on observed rewards and state transitions. The aim is for the agent to learn an optimal policy specifying the best action in each state for maximum long-term reward. Notable

TABLE 5. RL-based XRS research works.

RL-based approach	Algorithmic Approach	Key XRS research works mentioned
Value-based methods	DQN	[93], [94], [95], [96]
Policy-based methods	REINFORCE, PPO	[35], [97], [98], [99], [100], [101], [102]
Hybrid methods	Actor-Critic	[36], [103], [104], [105]

works leveraging DQN include DRN [93], SADQN [94], UQDN [95], and FeedRec [96]. These studies utilize DQN and its variants to enhance the effectiveness of RS.

2) POLICY-BASED METHODS

Policy-based RL directly learns a policy, mapping states to actions without explicitly computing value functions. Instead of estimating action values, these methods parametrize and optimize the policy to maximize expected cumulative rewards. Policy-based approaches are preferred for continuous or high-dimensional action spaces where computing action values are impractical. Examples include REINFORCE [35], PPO [97], and Trust Region Policy Optimization (TRPO) [98]. Notable works in this approach utilize REINFORCE and its variants, including PGCR [99], PPG [100], and RML [101], while optimization techniques in DPG-FBE [102].

3) HYBRID METHODS

Hybrid-based RL methods blend value-based and policy-based approaches, aiming for enhanced performance. They leverage both a value function to estimate action quality and a policy to guide action selection. This combination enhances sample efficiency and stability. Actor-critic algorithms [36] exemplify hybrid methods, using separate networks for policy and value function representation. Notable works utilize Actor-Critic and its variants to enhance effectiveness in XR, including LIRD [103], DeepPage [104], and SLC [105].

Table 5 summarizes the RL-based research works mentioned in this review paper. This subsection discusses various RL-based implementation approaches comprising value-based, policy-based, and hybrid methods.

Value-based RL learns the value function to estimate the expected cumulative reward in a state, often employing algorithms like DQN to optimize the action-value function. In contrast, policy-based RL directly learns the policy mapping states to actions, which is suitable for continuous or stochastic environments. Hybrid RL methods combine aspects of both approaches, integrating value function estimation with policy optimization to enhance performance and stability in learning.

RQ4: How does RL contribute to achieving explainability within RS?

RL enhances explainability in RS by providing a transparent framework for decision-making. It learns optimal recommendation strategies through user interactions, generating trajectories of state-action pairs that reveal the reasoning behind recommendations. Stakeholders can understand why certain recommendations are prioritized by analyzing reward signals and inspecting model parameters, enhancing overall system explainability. Despite RL's contributions to XRS development through value-based, policy-based, and hybrid approaches, challenges persist in achieving robust model-intrinsic explainability for generated recommendations.

E. INTEGRATED APPROACH OF KG AND RL-based METHODS

KG enriches RSs with semantic data, aiding in better user-item reasoning. Integrating KGs into RL enables informed item recommendations by considering user preferences and contextual factors. This section outlines KG and RL-based approaches for explainability in RSs, addressing RQ5. These methods are categorized into KG-RL-value-based, KG-RL-policy-based, and KG-RL-hybrid streams, as shown in Fig. 8.

1) KG-RL-VALUE-BASED METHODS

KG-RL-Value-based methods improve XRSs by combining a value-based RL framework, such as DQN, with KGs to enhance the explainability of recommendations. Notable works in this area include DPAO [39], RMS [41], KGQR [46], KGDQN [54], CDARL [64], and KPRLN [60], as listed in Table 3.

Jung et al. [39] introduce DPAO, employing dual policy learning with two deep Q-network models to determine high-order connectivity for aggregating users and items following an MDP. Ning et al. [41] propose RMS, an RL-based meta-path selection framework, utilizing a policy network to identify high-quality meta-paths and constructing a policy π maximizing the reward through an MDP and DQN. Zhou et al. [46] present the KGQR model, utilizing GCN to transform KG information, generating state representation, and recommending the highest-scored item from the candidate set through the Q-network. Xu et al. [54] propose KGDQN, combining KG and RL with a TransE embedding algorithm for KG and multi-layer Deep Q-learning for RL to predict rewards from reasoning paths. Wu et al. [60] propose KPRLN, extracting fine-grained user preference features from the KG and constructing a path network of user historical interaction items in the KG using deep RL, exploring the KG via cluster expansion with a double DQN model, and designing an attention mechanism to propagate higher-order user interests and aggregate user and item representations for prediction. These studies illustrate how KG integration in the value-based RL framework enhances recommendation explainability.

2) KG-RL-POLICY-BASED METHODS

KG-RL-Policy-based methods enhance XRSs by integrating policy-based RL algorithms like REINFORCE and policy optimization with KGs to improve recommendation explainability. Unlike DQN, these methods directly optimize the policy π. Key works in this area include PGPR [44], UPGPR [37], IP-PGPR [56], MES [55], SAPL [45], CERec [47], REKS [48], TMER-RL [49], PPO [51], Mcore [52], TN [57], RSL-GRU [58], HINRL4Rec [62], and HR-RL-KG [65], as listed in Table 3.

Xian et al. [44] introduce PGPR, integrating recommendation and interpretability through KGs. It formulates recommendation as a deterministic MDP over the KG utilizing RL with the REINFORCE algorithm and features, like soft reward strategy, user-conditional action pruning, and a multi-hop scoring function. The soft reward strategy employs a multi-hop scoring function utilizing heterogeneous KG information. Frej et al. [37] extend PGPR to UPGPR, allowing path patterns of any type and length and introducing a new reward mechanism to broaden exploration and enhance generalizability. Balloccu et al. [56] extend PGPR with in-processing and post-processing optimization techniques, generating substantially higher-quality reasoning paths. Tiwary et al. [55] extend PGPR to evaluate the explainability of recommendations with MES, a max explainability score based on rules, probability, entropy value, and reward from chosen traversal paths. Park et al. [45] propose SAPL, learning an RL policy to generate item recommendations and reasoning, enriching the KG with sentiment-related labels. Wang et al. [47] propose CERec, a counterfactual explainable recommendation method, generating item attribute-based counterfactual explanations, reducing the search space with an adaptive path sampler, and enhancing recommendation with the explanation policy. Vo [62] propose HINRL4Rec, integrating heterogeneous network embedding with a policy-guided path-based searching mechanism using RL. Other similar approaches adopting the REINFORCE algorithm with MDP include REKS [48], TMER-RL [49], Mcore [52], TN [57], RSL-GRU [58], and HR-RL-KG [65]. A few models adopted different policy optimization algorithms, such as PPO [97] and TRPO [98]. Feng et al. [51] introduce a proximal policy optimization XR model and utilize PPO to optimize recommendations. These studies illustrate how KG integration in the policy-based RL framework enhances recommendation explainability.

3) KG-RL-HYBRID-BASED METHODS

KG-RL-Hybrid-based methods improve XRSs by combining a hybrid-based RL framework, such as actor-critic, with KGs to enhance the explainability of recommendations. The actor-critic method combines value-based and policy-based methods by using two different networks, where an actor network uses a policy-based, and the critic uses a value-based to evaluate the policy learned by the agent. Prominent works in this area include ReMR [40], AnchorKG [42], ADAC [43], KGRL [53], KRRL [59], MEUR [63], RCENR [38], RKGR-RNS [50], and IMRL [61], as listed in Table 3.

Wang et al. [40] introduce ReMR, a model for multi-level recommendation reasoning over KGs, leveraging both ontology-view and instance-view KGs to capture multi-level user interests. They employ abstract MDP using cascading Actor-Critic for multi-level reasoning, adopting a top-down strategy to prune the search space and ensure that high-level KG knowledge guides low-level reasoning policies to a satisfying solution. Liu et al. [42] propose AnchorKG, which generates a compact anchor KG to enhance article latent representation. They frame anchor graph generation as a deterministic MDP and use RL techniques, specifically the actor-critic algorithm, for optimization. Zhao et al. [43] propose ADAC, leveraging human demonstrations to improve path searching. ADAC adopts adversarial imitation learning and expert paths to facilitate the search process for optimal paths in the KG. Chen et al. [53] introduce KGRL, utilizing KGs to enhance the critic's evaluation of generated policies. Unlike previous studies, KGRL embeds the KG into the actor-critic network, guiding the actor to find better recommendation policies by measuring proximity from the optimal path. Lin et al. [59] propose KRRL, a multi-level representation learning method enhancing semantic representation and relations in the KG. They adopt an RL approach, specifically the actor-critic algorithm, to guide path reasoning. Other similar approaches adopting the actor-critic algorithm with MDP include MEUR [63], RCENR [38], RKGR-RNS [50], and IMRL [61]. These studies illustrate how KG integration in the hybrid-based RL framework enhances recommendation explainability.

TABLE 6. KG-RL-based XRS research works.

KG-RL-based approach	Key XRS research works mentioned in this review paper
KG-RL-Value-based methods	[39], [41], [46], [54], [60], [64]
KG-RL-Policy-based methods	[37], [38], [44], [45], [47], [48], [49], [51], [52], [55], [56], [57], [58], [62], [65], [106]
KG-RL-Hybrid methods	[38], [40], [42], [43], [50], [53], [59], [61], [63]

Table 6 summarizes the KG-RL-based research works listed in this review paper. This subsection discusses various RL-based implementation approaches comprising value-based, policy-based, and hybrid.

KG-RL-Value-based methods focus on learning the value function and estimating the expected cumulative reward of actions in a given state under the KG. Algorithms like DQN are commonly employed to determine the optimal action-value function. On the other hand, KG-RL-Policy-based methods directly learn the policy mapping states to actions without explicitly computing a value function with the underlying KG. This approach is beneficial for continuous action spaces and stochastic environments. KG-RL-Hybrid methods combine elements of value-based and policy-based approaches along with KG, aiming to leverage their respective strengths. They often incorporate value function estimation alongside policy optimization techniques to achieve better performance and stability in learning.

RQ5: How important is it to integrate KG and RL for enhancing the effectiveness of the XRS?

KG and RL are indispensable in developing effective XRS. KG acts as a knowledge base, enabling the generation of explanations or reasoning for recommendations. RL, on the other hand, helps create a robust ML algorithm that captures user behaviour. Integrating these techniques results in a robust XRS, capturing embedded reasoning in model decisions and enabling effective traversal over the KG for model-intrinsic explainability. The research works discussed in section VI-E underscore the significance of integrating KG and RL in XRS.

F. KEY INSIGHTS

This section explores insights and discusses influential research in this domain. It highlights a few case studies demonstrating the advantages of KG and RL integration in XRS alongside the challenges associated with implementing such models in practice.

1) INFLUENTIAL RESEARCH

The review paper emphasizes several influential research papers that utilize both KG and RL in the context of XRS. Specifically, it mentions PGPR [44], KGQR [46], ReMR [40], and MES [55] as studies that continue to explore various RL approaches applied to KGs.

Xian et al. [44] introduced PGPR, which combined MDP and RL to navigate multi-hop paths in a KG for making recommendations. This research was pivotal as it was the first published research to apply RL over KG for developing RSs. It used the policy-based RL algorithm, REINFORCE, and formulated recommendations as a deterministic MDP over the KG to identify the optimal recommendation policy. It adopted a soft reward strategy and employed a multi-hop scoring function that utilized heterogeneous KG information. The approach faced challenges related to computational costs, personalization, optimal product suggestions, and the associated reasoning. The approach prioritized reasoning path selection based solely on path probability, overlooking other factors.

Later, research began to explore the effectiveness of value-based RL algorithms. Zhou et al. [46] developed the KGQR model, which replaced policy-based RL with value-based methods and implemented an RS using deep Q-networks. They leveraged prior knowledge of item correlations learned from the KG to guide candidate selection for improved item retrieval. They enhanced item and user state representations, propagated user preferences across correlated items in the KG to address user feedback sparsity and utilized GCN to transform KG information and generate state representations. The model recommended the items by selecting the highest-scored item from the candidate set through the Q-network.

There is a growing need to combine the advantages of both policy-based and value-based RL methods. Wang et al. [40] introduced ReMR, a model designed for multi-level recommendation reasoning across KGs. This model leverages both ontology-view and instance-view KGs to capture diverse user interests at multiple levels, thereby enhancing the accuracy of recommendations. The approach integrates RL methods, specifically employing a cascading Actor-Critic framework within an abstract MDP. This enables effective multi-level reasoning by applying a top-down strategy to streamline the search process, ensuring that higher-level KG knowledge guides lower-level reasoning policies toward optimal solutions. Additionally, the framework includes a method for extracting multi-level reasoning paths, automatically selecting between high-level concepts and low-level details to construct paths that effectively reveal user interests.

Recently, Tiwary et al. [55] expanded their research by focusing on explainability metrics to assess the clarity of generated recommendations. They introduced MES, a metric designed to evaluate the effectiveness of explainability in their recommendations by providing the max explainability score. This metric utilizes fundamental evaluation parameters within the RL framework, including the number of hops the agent takes to generate recommendations, the traversal path's transition probabilities, associated entropy, and rewards received during the path. These foundational parameters are leveraged to derive an evaluation metric that quantifies the quality of explainability provided by the recommendations.

2) CASE STUDIES

The integration of KG and RL contributes significantly to enhancing the quality of generated recommendations and facilitates achieving explainability for these recommendations. Research combining KG and RL has shown that integrated models are more effective in generating recommendations than those relying solely on KG. For instance, Wang et al. [40] demonstrated the superiority of his developed model, ReMR, by comparing it to various baseline models that utilized different variations of KG. His results indicated that ReMR outperformed the baseline models. Similar findings have been observed in other experiments where RL variations were applied to KGs. The experiments have demonstrated consistent performance across various datasets from different domains, such as movies (MovieLens, DoubanMovie), e-commerce (Amazon - Book, Clothing, CDs & Vinyl, Cellphones, Beauty), and music (Last.fm, KKBox), among others.

3) CHALLENGES WITH APPLYING KG AND RL IN XRS

Applying KG and RL in XRS presents intricate challenges. One significant hurdle is ensuring the interpretability of the knowledge representation derived from KGs. While KGs offer a structured framework of entities and their relationships, the decision-making processes of RL models trained on these graphs can be complex and opaque to end-users. This complexity undermines the system's ability to provide clear explanations for recommendations, potentially leading to user mistrust or dissatisfaction. Additionally, handling uncertainties and incompleteness within KGs poses another critical challenge. RL algorithms must navigate these nuances effectively to generate reliable recommendations while transparently communicating the limitations or uncertainties to users. Moreover, the dynamic nature of KGs requires continuous adaptation of RL models, necessitating robust mechanisms to update and maintain the knowledge base without compromising system performance or interpretability. Addressing these challenges requires advancements in AI algorithms and user-centric design to ensure that XRS effectively balances complexity with transparency, fostering user trust and acceptance.

G. EVALUATION METRICS

The effectiveness of the XRS model's recommendations is evaluated using four distinct metrics: Normalized Discounted Cumulative Gain (NDCG), Hit Rate (HR), Recall, and Precision. These metrics provide quantitative measures to assess the model's performance and effectiveness. NDCG captures the relevance and ranking of recommended items, HR quantifies the model's ability to suggest relevant items, Recall measures the model's capability to retrieve relevant items, and Precision gauges the accuracy of the recommended items. This range of evaluation metrics aims to provide a comprehensive and multifaceted assessment of the proposed XRS model's performance.

The quantitative evaluation of explainability is an advanced area, and various research efforts, including MES [55], are ongoing.

H. FUTURE WORKS

Future research may focus on making KG-RL-based RSs more interpretable by developing clear explanations for recommendations derived from complex KG interactions. Integrating diverse data types into KGs will enrich recommendations while ensuring dynamic adaptability and addressing privacy concerns will be crucial. Given the promising nature of the field and the significant interest within the research community in responsible AI, it is anticipated that the field of XRS will continue to pose new challenges and inspire further research endeavours. This will be essential to meet the demand for AI systems that are ethical, trustworthy, responsible, reliable, fair, transparent, secure, accountable, and explainable. Establishing standardized evaluation metrics will further enhance the reliability and transparency of these systems across different domains.

VII. CONCLUSION

In conclusion, this review paper provides a comprehensive overview of the current landscape of XRS, emphasizing the critical role of explainability and the contributions of KG and RL techniques in advancing RSs. The paper begins by summarizing existing reviews on XRS, outlining their contributions and areas of focus. It then describes a tailored web scraping tool developed to efficiently gather and organize research papers from leading platforms such as IEEE, ScienceDirect, ACM, and Springer online databases, facilitating the compilation of relevant research works for analysis. Through a systematic examination of research questions, the paper explores the necessity of explainability in RSs. It delves into various implementation approaches leveraging KG and RL, drawing insights from relevant research papers. By analyzing these studies, the paper sheds light on the significance of integrating KG and RL techniques to enhance the effectiveness and interpretability of RSs. It identifies emerging trends in the field and recommends methods for developing XRS based on the insights gleaned from the reviewed studies.

Overall, the review paper offers valuable insights into the evolving landscape of XRS, highlighting the importance of explainability and the pivotal role played by KG and RL techniques.

DECLARATION OF CONFLICTING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

[1] A. Vultureanu-Albisi and C. Badica, "Recommender systems: An explainable AI perspective," in *Proc. Int. Conf. Innov. Intell. Syst. Appl. (INISTA)*, Kocaeli, Turkey, Aug. 2021, pp. 1–6, doi: 10.1109/INISTA52262.2021.9548125.

[2] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Int. J. Surg.*, vol. 8, no. 5, pp. 336–341, 2010.

[3] X. Chen, L. Yao, J. McAuley, G. Zhou, and X. Wang, "Deep reinforcement learning in recommender systems: A survey and new perspectives," *Knowl.-Based Syst.*, vol. 264, Mar. 2023, Art. no. 110335.

[4] Q. Guo, F. Zhuang, C. Qin, H. Zhu, X. Xie, H. Xiong, and Q. He, "A survey on knowledge graph-based recommender systems," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 8, pp. 3549–3568, Aug. 2022.

[5] I. Pesovski, A. M. Bogdanova, and V. Trajkovik, "Systematic review of the published explainable educational recommendation systems," in *Proc. 20th Int. Conf. Inf. Technol. Based Higher Educ. Training (ITHET)*, Antalya, Turkey, Nov. 2022, pp. 1–8, doi: 10.1109/ITHET56107.2022.10032029.

[6] N. A. Osman, S. A. Mohd Noah, M. Darwich, and M. Mohd, "Integrating contextual sentiment analysis in collaborative recommender systems," *PLoS ONE*, vol. 16, no. 3, Mar. 2021, Art. no. e0248695.

[7] S. M. Al-Ghuribi, S. A. M. Noah, M. A. Mohammed, S. N. Qasem, and B. A. H. Murshed, "To cluster or not to cluster: The impact of clustering on the performance of aspect-based collaborative filtering," *IEEE Access*, vol. 11, pp. 41979–41994, 2023.

[8] Y. Bekkemoen, "Explainable reinforcement learning (XRL): A systematic literature review and taxonomy," *Mach. Learn.*, vol. 113, no. 1, pp. 355–441, Jan. 2024.

[9] M. R. Naqvi, L. Elmhadhbi, A. Sarkar, B. Archimede, and M. H. Karray, "Survey on ontology-based explainable AI in manufacturing," *J. Intell. Manuf.*, Feb. 2024, doi: 10.1007/s10845-023-02304-z.

[10] R. Dazeley, P. Vamplew, and F. Cruz, "Explainable reinforcement learning for broad-XAI: A conceptual framework and survey," *Neural Comput. Appl.*, vol. 35, no. 23, pp. 16893–16916, Aug. 2023.

[11] W. Saeed and C. Omlin, "Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities," *Knowl.-Based Syst.*, vol. 263, Mar. 2023, Art. no. 110273.

[12] X.-H. Li, C. C. Cao, Y. Shi, W. Bai, H. Gao, L. Qiu, C. Wang, Y. Gao, S. Zhang, X. Xue, and L. Chen, "A survey of data-driven and knowledge-aware eXplainable AI," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 1, pp. 29–49, Jan. 2022.

[13] Y. Zhang and X. Chen, "Explainable recommendation: A survey and new perspectives," *Found. Trends Inf. Retr.*, vol. 14, no. 1, pp. 1–101, 2020.

[14] S. M. Al-Ghuribi and S. A. Mohd Noah, "Multi-criteria review-based recommender system—The state of the art," *IEEE Access*, vol. 7, pp. 169446–169468, 2019.

[15] R. M. Nawi, S. A. M. Noah, and L. Q. Zakaria, "Integration of linked open data in collaborative group recommender systems," *IEEE Access*, vol. 9, pp. 150753–150767, 2021.

[16] R. R. Hoffman, S. T. Mueller, G. Klein, and J. Litman, "Metrics for explainable AI: Challenges and prospects," 2018, *arXiv:1812.04608*.

[17] A. A. Khan, S. Badshah, P. Liang, M. Waseem, B. Khan, A. Ahmad, M. Fahmideh, M. Niazi, and M. A. Akbar, "Ethics of AI: A systematic literature review of principles and challenges," in *Proc. 26th Int. Conf. Eval. Assessment Softw. Eng. (EASE)*, Gothenburg, Sweden, 2022, pp. 383–392, doi: 10.1145/3530019.3531329.

[18] P. Voigt and A. Von dem Bussche, *The EU General Data Protection Regulation (GDPR)*, vol. 10, 1st ed., Cham, Switzerland: Springer, 2017, p. 383, doi: 10.1007/978-3-319-57959-7.

[19] J. Laux, S. Wachter, and B. Mittelstadt, "Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk," *Regulation Governance*, vol. 18, no. 1, pp. 3–32, 2023.

[20] L. de la Torre, "A guide to the California consumer privacy act of 2018," *Available at SSRN*, Nov. 2018, Art. no. 3275571.

[21] B. Kocak, A. H. Yardimci, S. Yuzkan, A. Keles, O. Altun, E. Bulut, O. N. Bayrak, and A. A. Okumus, "Transparency in artificial intelligence research: A systematic review of availability items related to open science in radiology and nuclear medicine," *Academic Radiol.*, vol. 30, no. 10, pp. 2254–2266, Oct. 2023.

[22] A. Hogan, E. Blomqvist, M. Cochez, C. d'Amato, G. D. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, S. Neumaier, and A. C. N. Ngomo, "Knowledge graphs," *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–37, 2021.

[23] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor, "Freebase: A collaboratively created graph database for structuring human knowledge," in *Proc. ACM SIGMOD Int. Conf. Manag. Data (SIGMOD)*, New York, NY, USA, 2008, pp. 1247–1250.

[24] J. Lehmann, R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P. N. Mendes, S. Hellmann, M. Morsey, P. Van Kleef, S. Auer, and C. Bizer, "DBpedia—A large-scale, multilingual knowledge base extracted from Wikipedia," *Semantic web*, vol. 6, no. 2, pp. 167–195, 2015.

[25] F. M. Suchanek, G. Kasneci, and G. Weikum, "Yago: A core of semantic knowledge," in *Proc. 16th Int. Conf. World Wide Web (WWW)*, New York, NY, USA, 2007, pp. 697–706, doi: 10.1145/1242572.1242667.

[26] A. Singhal. (2012). *Introducing the Knowledge Graph: Things, Not Strings*. [Online]. Available: https://blog.google/products/search/introducing-knowledge-graph-things-not/

[27] P. A. Bonatti, S. Decker, A. Polleres, and V. Presutti, "Knowledge graphs: New directions for knowledge representation on the semantic web (dagstuhl seminar 18371)," *Dagstuhl Rep.*, vol. 8, no. 9, pp. 29–111, 2019.

[28] M. Yahya, J. G. Breslin, and M. I. Ali, "Semantic web and knowledge graphs for industry 4.0," *Appl. Sci.*, vol. 11, no. 11, p. 5110, May 2021.

[29] D. Fensel, U. Şimşek, K. Angele, E. Huaman, E. Kärle, O. Panasiuk, I. Toma, J. Umbrich, and A. Wahler, *Knowledge Graphs*, 1 ed. Cham, Switzerland: Springer, 2020, p. 148, doi: https://doi.org/10.1007/978-3-030-37439-6

[30] S. J. Russell, *Artificial Intelligence: A Modern Approach*, 1 ed. London, U.K.: Pearson, 2010.

[31] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogamani, and P. Pérez, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 4909–4926, Jun. 2022.

[32] Z. Yan, J. Ge, Y. Wu, L. Li, and T. Li, "Automatic virtual network embedding: A deep reinforcement learning approach with graph convolutional networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 6, pp. 1040–1057, Jun. 2020.

[33] R. Wickman, "SPARRL: Graph sparsification via deep reinforcement," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, Philadelphia, PA, USA, vol. 2022, pp. 2521–2523.

[34] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, pp. 279–292, May 1992.

[35] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Mach. Learn.*, vol. 8, no. 3, pp. 229–256, 1992.

[36] V. Konda and J. Tsitsiklis, "Actor-critic algorithms," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 12, 1999, pp. 1–7.

[37] J. Frej, N. Shah, M. Knezevic, T. Nazaretsky, and T. Käser, "Finding paths for explainable MOOC recommendation: A learner perspective," in *Proc. 14th Learn. Anal. Knowl. Conf.*, Kyoto, Japan, 2024, pp. 426–437, doi: 10.1145/3636555.3636898.

[38] H. Jiang, C. Li, J. Cai, and J. Wang, "RCENR: A reinforced and contrastive heterogeneous network reasoning model for explainable news recommendation," in *Proc. 46th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Taiwan, 2023, pp. 1710–1720, doi: 10.1145/3539618.3591753.

[39] H. Jung, S. Kim, and H. Park, "Dual policy learning for aggregation optimization in graph neural network-based recommender systems," in *Proc. ACM Web Conf.*, Austin, TX, USA, 2023, pp. 1478–1488, doi: 10.1145/3543507.3583241

[40] X. Wang, K. Liu, D. Wang, L. Wu, Y. Fu, and X. Xie, "Multi-level recommendation reasoning over knowledge graphs with reinforcement learning," in *Proc. ACM Web Conf.*, Lyon, France, 022, pp. 2098–2108, doi: 10.1145/3485447.3512083

[41] W. Ning, R. Cheng, J. Shen, N. A. H. Haldar, B. Kao, X. Yan, N. Huo, W. K. Lam, T. Li, and B. Tang, "Automatic meta-path discovery for effective graph-based recommendation," in *Proc. 31st ACM Int. Conf. Inf. Knowl. Manag.*, Atlanta, GA, USA, 2022, pp. 1563–1572, doi: 10.1145/3511808.3557244.

[42] D. Liu, J. Lian, Z. Liu, X. Wang, G. Sun, and X. Xie, "Reinforced anchor knowledge graph generation for news recommendation reasoning," in *Proc. 27th ACM SIGKDD Conf. Knowl. Discovery Data Mining (KDD)*, Singapore, vol. 2021, pp. 1055–1065.

[43] K. Zhao, X. Wang, Y. Zhang, L. Zhao, Z. Liu, C. Xing, and X. Xie, "Leveraging demonstrations for reinforcement recommendation reasoning over knowledge graphs," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Beijing, China, 2020, pp. 239–248, doi: 10.1145/3397271.3401171.

[44] Y. Xian, Z. Fu, G. De Melo, and Y. Zhang, "Reinforcement knowledge graph reasoning for explainable recommendation," in *Proc. 42nd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, Paris, France, 2019, pp. 285–294, doi: 10.1145/3331184.3331203.

[45] S.-J. Park, D.-K. Chae, H.-K. Bae, S. Park, and S.-W. Kim, "Reinforcement learning over sentiment-augmented knowledge graphs towards accurate and explainable recommendation," in *Proc. 15th ACM Int. Conf. Web Search Data Mining*, Phoenix, AZ, USA, 2022, pp. 784–793, doi: 10.1145/3488560.3498515.

[46] S. Zhou, X. Dai, H. Chen, W. Zhang, K. Ren, R. Tang, X. He, and Y. Yu, "Interactive recommender system via knowledge graph-enhanced reinforcement learning," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, New York, NY, USA, 2020, pp. 179–188, doi: 10.1145/3397271.3401174.

[47] X. Wang, Q. Li, D. Yu, Q. Li, and G. Xu, "Reinforced path reasoning for counterfactual explainable recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 7, pp. 3443–3459, Jul. 2024, doi: 10.1109/TKDE.2024.3354077.

[48] H. Wu, H. Fang, Z. Sun, C. Geng, X. Kong, and Y.-S. Ong, "A generic reinforced explainable framework with knowledge graph for session-based recommendation," in *Proc. IEEE 39th Int. Conf. Data Eng. (ICDE)*, Anaheim, CA, USA, Apr. 2023, pp. 1260–1272, doi: 10.1109/ICDE55515.2023.00101.

[49] Y. Li, H. Chen, Y. Li, L. Li, P. S. Yu, and G. Xu, "Reinforcement learning based path exploration for sequential explainable recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 11, pp. 11801–11814, 2023.

[50] S. Zhang, Y. Ouyang, Z. Liu, W. Rong, and Z. Xiong, "Reinforcement learning-based explainable recommendation over knowledge graphs with negative sampling," in *Proc. IEEE Smartworld, Ubiquitous Intell. Comput., Scalable Comput. Commun., Digit. Twin, Privacy Comput., Metaverse, Auton. & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta)*, Haikou, China, Dec. 2022, pp. 1948–1953.

[51] Q. Feng, G. Xiao, Y. Liang, H. Zhang, L. Yan, and X. Yi, "Proximal policy optimization for explainable recommended systems," in *Proc. 4th Int. Conf. Data-Driven Optim. Complex Syst. (DOCS)*, Chengdu, China, Oct. 2022, pp. 1–6, doi: 10.1109/DOCS55193.2022.9967709.

[52] X. Li, Y. Shen, and L. Chen, "Mcore: Multi-agent collaborative learning for knowledge-graph-enhanced recommendation," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Auckland, New Zealand, Dec. 2021, pp. 330–339, doi: 10.1109/ICDM51629.2021.00044.

[53] X. Chen, C. Huang, L. Yao, X. Wang, W. Liu, and W. Zhang, "Knowledge-guided deep reinforcement learning for interactive recommendation," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Glasgow, U.K., Jul. 2020, pp. 1–8, doi: 10.1109/IJCNN48605.2020.9207010.

[54] W. Xu, X. Gao, Y. Sheng, and G. Chen, "Recommendation system with reasoning path based on DQN and knowledge graph," in *Proc. 15th Int. Conf. Ubiquitous Inf. Manage. Commun. (IMCOM)*, Seoul, South Korea, Jan. 2021, pp. 1–8, doi: 10.1109/IMCOM51814.2021.9377414.

[55] N. Tiwary, S. A. Mohd Noah, F. Fauzi, and T. S. Yee, "Max explainability score—A quantitative metric for explainability evaluation in knowledge graph-based recommendations," *Comput. Elect. Eng.*, vol. 116, May 2024, Art. no. 109190.

[56] G. Balloccu, L. Boratto, G. Fenu, and M. Marras, "Reinforcement recommendation reasoning through knowledge graphs for explanation path quality," *Knowl.-Based Syst.*, vol. 260, Jan. 2023, Art. no. 110098.

[57] M. Saebi, S. Kreig, C. Zhang, M. Jiang, T. Kajdanowicz, and N. V. Chawla, "Heterogeneous relational reasoning in knowledge graphs with reinforcement learning," *Inf. Fusion*, vol. 88, pp. 12–21, Dec. 2022.

[58] Z. Cui, H. Chen, L. Cui, S. Liu, X. Liu, G. Xu, and H. Yin, "Reinforced KGs reasoning for explainable sequential recommendation," *World Wide Web*, vol. 25, no. 2, pp. 631–654, Mar. 2022.

[59] Y. Lin, W. Zhang, F. Lin, W. Zeng, X. Zhou, and P. Wu, "Knowledge-aware reasoning with self-supervised reinforcement learning for explainable recommendation in MOOCs," *Neural Comput. Appl.*, vol. 36, no. 8, pp. 4115–4132, Mar. 2024.

[60] D. Wu, M. Tang, S. Zhang, A. You, and W. Gao, "KPRLN: Deep knowledge preference-aware reinforcement learning network for recommendation," *Complex Intell. Syst.*, vol. 9, no. 6, pp. 6645–6659, Dec. 2023.

[61] X. Chen, S. Wang, L. Qi, Y. Li, and L. Yao, "Intrinsically motivated reinforcement learning based recommendation with counterfactual data augmentation," *World Wide Web*, vol. 26, no. 5, pp. 3253–3274, Sep. 2023.

[62] T. Vo, "An integrated network embedding with reinforcement learning for explainable recommendation," *Soft Comput.*, vol. 26, no. 8, pp. 3757–3775, Apr. 2022.

[63] Z. Liang, L. Mu, J. Chen, and Q. Xie, "Graph path fusion and reinforcement reasoning for recommendation in MOOCs," *Educ. Inf. Technol.*, vol. 28, no. 1, pp. 525–545, Jan. 2023.

[64] Z. Liu, Y. Ma, M. Hildebrandt, Y. Ouyang, and Z. Xiong, "CDARL: A contrastive discriminator-augmented reinforcement learning framework for sequential recommendations," *Knowl. Inf. Syst.*, vol. 64, no. 8, pp. 2239–2265, 2022.

[65] W. Song, T. Wang, and Z. Zhang, "Recommendations based on reinforcement learning and knowledge graph," in *Proc. IEA/AIE: Adv. Trends Artif. Intell. Theory Appl.*, vol. 13925. Cham, Switzerland: Springer, 2023, pp. 313–324, doi: 10.1007/978-3-031-36819-6_28.

[66] Q. Ai, V. Azizi, X. Chen, and Y. Zhang, "Learning heterogeneous knowledge base embeddings for explainable recommendation," *Algorithms*, vol. 11, no. 9, p. 137, Sep. 2018.

[67] P. Valdiviezo-Diaz, F. Ortega, E. Cobos, and R. Lara-Cabrera, "A collaborative filtering approach based on Naïve Bayes classifier," *IEEE Access*, vol. 7, pp. 108581–108592, 2019.

[68] A. Jung and P. H. J. Nardelli, "An information-theoretic approach to personalized explainable machine learning," *IEEE Signal Process. Lett.*, vol. 27, pp. 825–829, 2020.

[69] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 26. South Lake Tahoe, CA, USA: Curran Associates 2013, pp. 2787–2795.

[70] Z. Wang, J. Zhang, J. Feng, and Z. Chen, "Knowledge graph embedding by translating on hyperplanes," in *Proc. 28th AAAI Conf. Artif. Intell. (AAAI)*, Québec City, QC, Canada, vol. 28, 2014, pp. 1112–1119, doi: 10.1609/aaai.v28i1.8870.

[71] Y. Lin, Z. Liu, M. Sun, Y. Liu, and X. Zhu, "Learning entity and relation embeddings for knowledge graph completion," in *Proc. AAAI Conf. Artif. Intell.*, vol. 29, Austin, TX, USA, 2015, pp. 2181–2187, doi: 10.1609/aaai.v29i1.9491.

[72] C. Xie, "Commodity knowledge graph-based TransD-KGAT method for recommendation," in *Proc. 5th Int. Conf. Comput. Inf. Sci. Artif. Intell. (CISAI)*, vol. 12566, Chongqing, China. Bellingham, WA, USA: SPIE, 2023, Art. no. 125662F, doi: 10.1117/12.2667700.

[73] B. Yang, W. Yih, X. He, J. Gao, and L. Deng, "Embedding entities and relations for learning and inference in knowledge bases," 2014, *arXiv:1412.6575*.

[74] F. Zhang, N. J. Yuan, D. Lian, X. Xie, and W.-Y. Ma, "Collaborative knowledge base embedding for recommender systems," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2016, pp. 353–362, doi: 10.1145/2939672.2939673.

[75] J. Huang, W. X. Zhao, H. Dou, J.-R. Wen, and E. Y. Chang, "Improving sequential recommendation with knowledge-enhanced memory networks," in *Proc. 41st Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, New York, NY, USA, 2018, pp. 505–514, doi: 10.1145/3209978.3210017.

[76] D. Yang, Z. Guo, Z. Wang, J. Jiang, Y. Xiao, and W. Wang, "A knowledge-enhanced deep recommendation framework incorporating GAN-based models," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Singapore, Nov. 2018, pp. 1368–1373, doi: 10.1109/ICDM.2018.00187.

[77] Y. Ye, X. Wang, J. Yao, K. Jia, J. Zhou, Y. Xiao, and H. Yang, "Bayes EMbedding (BEM) refining representation by integrating knowledge graphs and behavior-specific networks," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manag. (CIKM)*, New York, NY, USA, 2019, pp. 679–688, doi: 10.1145/3357384.3358014.

[78] X. Yu, X. Ren, Q. Gu, Y. Sun, and J. Han, "Collaborative filtering with entity similarity regularization in heterogeneous information networks," in *Proc. IJCAI HINA*, vol. 27, 2013, pp. 1–6.

[79] C. Luo, W. Pang, Z. Wang, and C. Lin, "Hete-CF: Social-based collaborative filtering recommendation using heterogeneous relations," in *Proc. IEEE Int. Conf. Data Mining*, Shenzhen, China, Dec. 2014, pp. 917–922, doi: 10.1109/ICDM.2014.64.

[80] X. Yu, X. Ren, Y. Sun, Q. Gu, B. Sturt, U. Khandelwal, B. Norick, and J. Han, "Personalized entity recommendation: A heterogeneous information network approach," in *Proc. 7th ACM Int. Conf. Web Search Data Mining*, New York, NY, USA, 2014, pp. 283–292, doi: 10.1145/2556195.2556259.

[81] B. Hu, C. Shi, W. X. Zhao, and P. S. Yu, "Leveraging meta-path based context for top-N recommendation with a neural co-attention model," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2018, pp. 1531–1540, doi: 10.1145/3219819.3219965.

[82] Z. Sun, J. Yang, J. Zhang, A. Bozzon, L.-K. Huang, and C. Xu, "Recurrent knowledge graph embedding for effective recommendation," in *Proc. 12th ACM Conf. Recommender Syst.*, New York, NY, USA, 2018, pp. 297–305, doi: 10.1145/3240323.3240361.

[83] H. Zhao, Q. Yao, J. Li, Y. Song, and D. L. Lee, "Meta-graph based recommendation fusion over heterogeneous information networks," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2017, pp. 635–644, doi: 10.1145/3097983.3098063.

[84] X. Wang, D. Wang, C. Xu, X. He, Y. Cao, and T.-S. Chua, "Explainable reasoning over knowledge graphs for recommendation," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, 2019, pp. 5329–5336, doi: 10.1609/aaai.v33i01.33015329.

[85] H. Wang, F. Zhang, J. Wang, M. Zhao, W. Li, X. Xie, and M. Guo, "RippleNet: Propagating user preferences on the knowledge graph for recommender systems," in *Proc. 27th ACM Int. Conf. Inf. Knowl. Manag.*, New York, NY, USA, 2018, pp. 417–426, doi: 10.1145/3269206.3271739.

[86] X. Tang, T. Wang, H. Yang, and H. Song, "AKUPM: Attention-enhanced knowledge-aware user preference model for recommendation," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2019, pp. 1891–1899, doi: 10.1145/3292500.3330705.

[87] Q. Li, X. Tang, T. Wang, H. Yang, and H. Song, "Unifying task-oriented knowledge graph learning and recommendation," *IEEE Access*, vol. 7, pp. 115816–115828, 2019.

[88] H. Wang, M. Zhao, X. Xie, W. Li, and M. Guo, "Knowledge graph convolutional networks for recommender systems," in *Proc. World Wide Web Conf.*, New York, NY, USA, 2019, pp. 3307–3313, doi: 10.1145/3308558.3313417

[89] X. Wang, X. He, Y. Cao, M. Liu, and T.-S. Chua, "KGAT: Knowledge graph attention network for recommendation," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2019, pp. 950–958, doi: 10.1145/3292500.3330989.

[90] J. Zhao, Z. Zhou, Z. Guan, W. Zhao, W. Ning, G. Qiu, and X. He, "IntentGC: A scalable graph convolution framework fusing heterogeneous information for recommendation," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2019, pp. 2347–2357, doi: 10.1145/3292500.3330686.

[91] X. Sha, Z. Sun, and J. Zhang, "Hierarchical attentive knowledge graph embedding for personalized recommendation," *Electron. Commerce Res. Appl.*, vol. 48, Jul. 2021, Art. no. 101071.

[92] R. Shimizu, M. Matsutani, and M. Goto, "An explainable recommendation framework based on an improved knowledge graph attention network with massive volumes of side information," *Knowl.-Based Syst.*, vol. 239, Mar. 2022, Art. no. 107970.

[93] G. Zheng, F. Zhang, Z. Zheng, Y. Xiang, N. J. Yuan, X. Xie, and Z. Li, "DRN: A deep reinforcement learning framework for news recommendation," in *Proc. World Wide Web Conf.*, Geneva, Switzerland, 2018, pp. 167–176, doi: 10.1145/3178876.3185994

[94] Y. Lei, Z. Wang, W. Li, and H. Pei, "Social attentive deep Q-network for recommendation," in *Proc. 42nd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, New York, NY, USA, 2019, pp. 1189–1192, doi: 10.1145/3331184.3331302.

[95] Y. Lei and W. Li, "Interactive recommendation with user-specific deep reinforcement learning," *ACM Trans. Knowl. Discovery Data*, vol. 13, no. 6, pp. 1–15, Dec. 2019.

[96] L. Zou, L. Xia, Z. Ding, J. Song, W. Liu, and D. Yin, "Reinforcement learning to optimize long-term user engagement in recommender systems," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2019, pp. 2810–2818, doi: 10.1145/3292500.3330668.

[97] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017, *arXiv:1707.06347*.

[98] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *Proc. 32nd Int. Conf. Mach. Learn. (ICML)*, Lille, France, vol. 37, 2015, pp. 1889–1897. [Online]. Available: https://dl.acm.org/doi/abs/10.5555/3045118.3045319

[99] F. Pan, Q. Cai, P. Tang, F. Zhuang, and Q. He, "Policy gradients for contextual recommendations," in *Proc. World Wide Web Conf.*, New York, NY, USA, 2019, pp. 1421–1431, doi: 10.1145/3308558.3313616

[100] J. Xu, Z. Wei, L. Xia, Y. Lan, D. Yin, X. Cheng, and J.-R. Wen, "Reinforcement learning to rank with pairwise policy gradient," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, New York, NY, USA, 2020, pp. 509–518, doi: 10.1145/3397271.3401148.

[101] A. Montazeralghaem, H. Zamani, and J. Allan, "A reinforcement learning framework for relevance feedback," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, New York, NY, USA, 2020, pp. 59–68, doi: 10.1145/3397271.3401099.

[102] Y. Hu, Q. Da, A. Zeng, Y. Yu, and Y. Xu, "Reinforcement learning to rank in e-commerce search engine: Formalization, analysis, and application," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, Association for Computing Machinery, 2018, pp. 368–377, doi: 10.1145/3219819.3219846.

[103] X. Zhao, L. Zhang, L. Xia, Z. Ding, D. Yin, and J. Tang, "Deep reinforcement learning for list-wise recommendations," 2017, *arXiv:1801.00209*.

[104] X. Zhao, L. Xia, L. Zhang, Z. Ding, D. Yin, and J. Tang, "Deep reinforcement learning for page-wise recommendations," in *Proc. 12th ACM Conf. Recommender Syst.*, New York, NY, USA, 2018, pp. 95–103, doi: 10.1145/3240323.3240374.

[105] F. Liu, R. Tang, H. Guo, X. Li, Y. Ye, and X. He, "Top-aware reinforcement learning based recommendation," *Neurocomputing*, vol. 417, pp. 255–269, Dec. 2020.

[106] W. Song, Z. Duan, Z. Yang, H. Zhu, M. Zhang, and J. Tang, "Ekar: An explainable method for knowledge aware recommendation," 2019, *arXiv:1906.09506*.

NEERAJ TIWARY received the master's degree in computer science from the National Institute of Technology at Rourkela, India, in 2003. He is currently pursuing the Ph.D. degree in explainable artificial intelligence with Universiti Kebangsaan Malaysia (UKM). He received the Certificate in business analytics from Indian School of Business, Hyderabad, India, in 2016. He handled responsibilities as the Head of Data Science, AICOE Capability, and Recruitment Lead in a customer-focused role for AI and ML. His research interests include AI, including gen AI, KG, RL, ML, CV, text analytics, and language models.

SHAHRUL AZMAN MOHD NOAH (Member, IEEE) received the M.Sc. and Ph.D. degrees in information studies from The University of Sheffield. He is currently a Professor with the Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, and also heads the Text Analytics Research Group. His current research interests include semantic computing with special emphasis on information retrieval, ontology, and recommender systems. He has published various research articles in these areas. He is a member of the IEEE Computer Society and the International Association for Ontology and its Applications (IAOA).

FARIZA FAUZI received the Ph.D. degree in information technology from Monash University, in 2012. She is currently a Senior Lecturer with the Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia. Her research interests include web information extraction and processing, natural language processing, social media, and cybersecurity.

TAN SIOK YEE received the Ph.D. degree in industrial computing from Unversiti Kebangsaan Malaysia, in 2017. In 2018, she joined Universiti Kebangsaan Malaysia, where she is currently a Senior Lecturer and simultaneously as a Researcher with the Mixed Reality and Pervasive Laboratory (MyXLab). Her research interests include technology and innovation, with a particular focus on augmented reality, virtual reality, image processing, and mobile application development. She received several important recognitions for her research career, including the Fundamental Research Grant Scheme (FRGS), Translational Grant UKM (TR-UKM), and Research University Grant (GUP).