IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 1 February 2024, accepted 17 February 2024, date of publication 20 February 2024, date of current version 26 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3367971

RESEARCH ARTICLE

A Network Analysis-Driven Framework for Factual Explainability of Knowledge Graphs

SIRAJ MUNIR[1], RAUF AHMED SHAMS MALICK[2], AND STEFANO FERRETTI[1]
[1]Department of Pure and Applied Sciences, University of Urbino Carlo Bo, 61029 Urbino, Italy
[2]Department of Computer Science, National University of Computing and Emerging Sciences, Karachi 74200, Pakistan

Corresponding author: Siraj Munir (s.munir@campus.uniurb.it)

This work was supported in part by the National Operational Program (PON) Program, and in part by the University of Urbino Carlo Bo under Grant 317840.

ABSTRACT Knowledge Graphs are widely used to represent knowledge structures in complex domains. In most real-world scenarios, these knowledge structures are dynamic. As a result, measures must be developed to assess the robustness and usability of Knowledge Graphs in temporal settings. Additionally, the explainability of inherent knowledge constituents is crucial for the desired attention of Knowledge Graphs, particularly in temporal settings. In this paper, we developed a framework to understand the robustness of factual explainability of Knowledge Graphs. The method is further verified by using meso-level attributes of the knowledge graph. The complex network analysis along with the community structures are co-evaluated through homophilic and heterophilic properties within the graph to validate the robustness of the factual interpretations. The analysis reveals that symbolic representation could be used as a reasonable metric for extracting link-based communities.

INDEX TERMS Knowledge graph analysis, complex network analysis, homophily and heterophily analysis, factual reasoning and explainability, interpretability.

I. INTRODUCTION

With the advent of artificial intelligent applications, we have witnessed groundbreaking marvels. These applications entertain millions of users daily, contributing to the greater good. From virtual-assisted online shopping to movie recommendations and healthcare, we observe AI-assisted applications in action. While we leverage these systems for mutual benefit, there are areas, such as healthcare, military, and similar decision support systems, where accountability and responsibility are crucial in case of any issues. This is especially pertinent following the implementation of regulations like the European Data Protection Regulation (GDPR)[1] and others. Explainable AI (XAI) is a hot research area addressing the black-box nature of AI systems in a human-understandable and interpretable way [1]. However, despite significant progress, there are still open challenges that demand attention.

A growing trend in graph-based intelligent systems has recently been observed [2], [3], [4], [5]. Various solutions leveraging Graph Neural Networks (GNNs) have been proposed, given their effectiveness in handling structures represented as graphs [6]. GNNs excel in reasoning as they represent entities connected via edges, such as chemical compounds with bonds represented as graphs. However, a notable gap exists in the extensive testing of GNNs for semantic representations [7].

Similarly, social networks embody semantic relationships represented as interactions like "friend-of-friend" or "follower-followee." Several solutions have been proposed for learning and explaining these relationships [2]. However, there is room for improvement in reasoning methods for homophilic and heterophilic networks [2]. Homophily and heterophily analysis help us to identify and study the underlying attributes of community construction as it changes

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

[1]https://www.garanteprivacy.it/il-testo-del-regolamento

over time. Homophily represents a network's tendency for like-minded individuals to attract each other [8]. Conversely, heterophily represents the tendency of individuals to interact with diverse groups [9]. These notions are exploited to extract factual explanations or reasoning for KG.

To counter this gap we propose an end-to-end workflow that utilizes factual reasoning for exploiting the Knowledge Graph (KG) based representation(s) and Complex Network Analysis (NA) as a connecting block for analyzing networks statistically and structurally. Often these networks depict complex systems that exist in our daily lives (friendship networks, biological networks, etc.) [10], [11]. NA is a tool used to analyze networks through statistical measures like community detection, centrality measures, connections, and ties [12]. Factual reasoning is a special type based on evidence known as facts [13]. These facts are usually represented by formal logic (First Order Logic (FOL), Horn Clause, etc.). KGs efficiently integrate multi-dimensional information in a query-able and understandable manner, originally utilized by Google in 2012 [14] to enhance search results.

Network analysis (NA) can be divided into two main types: static and dynamic. Static NA examines a network at a single point in time, whereas dynamic NA tracks the changes in a network over time. In this study, we focus on dynamic NA, as KGs often exhibit temporal and structural complexity. To measure the properties and analyze selected KG, we use the following metrics: modularity, betweenness and closeness centrality, homophily, heterophily, and preferential attachment. Modularity quantifies the degree to which a network can be partitioned into communities or subgroups that reflect the overall system [15].

This study introduces a novel approach for knowledge extraction to factually explain the KG. This study is distinct from previous research by proposing an integrated methodology that incorporates symbolic rule extraction (facts), homophily and heterophily analysis, and factual reasoning for Knowledge Graphs (KG). The objectives for conducting this study are as follows:

- To study and analyze KGs and understand their features such as relationship dynamics, communities, and other properties, as a dynamic representation.
- To pipeline the methodology that processes the raw semantic triplets and extract reasonable (explainable or interpretable) facts for modeling KG.
- To utilize the KG for homophily and heterophily analysis to validate factual reasoning.
- To validate the proposed approach over real-world networks.
- Aggregating dynamic NA for observing facts at different snapshots.

We adopted a workflow with the following features to counter the objectives mentioned.

- An end-to-end workflow is presented that is capable of processing raw data. Extracts symbolic facts and rules (FOL) to represent decision boundary and models the KG.
- Incorporating stochastic models for homophily and heterophily analysis to confirm factual explanations or interpretations.
- A dynamic NA driven by interaction is presented to observe concrete facts within KG at different snapshots.
- For validation, we pilot the presented approach with a real-world dataset.

This study is organized into six sections for a detailed discussion of these contributions. Section I introduces the foundational aspects of this study. Section II discusses the current state of the art and related technologies. Section III delves into the methodological and experimental details of the proposed workflow. Section IV provides a brief discussion of the results we acquired from KG analysis specifically homophily and heterophily analysis. Section V highlights future avenues and limitations of this work. Finally, section VI concludes the study.

II. BACKGROUND AND RELATED TECHNOLOGIES

In this section, we will delve into the current state of the art. To begin our discussion, we will explore the background and related technologies that have paved the way for our proposed study.

A. KNOWLEDGE GRAPH AND SYMBOLIC REPRESENTATION

A Knowledge Graph (KG) is a highly effective semantic method for information representation in a query-able form. Everything in a KG is connected via relationships, typically representing connections between two entities. More formally, a KG is a graph G with nodes and edges (V and E), where V denotes entity nodes (such as Authors, Country, Seniority, etc.) and E represents relationships (Met, $Has\_Country$, $Has\_Attended$, etc.). Each node and edge connection can be expressed as a semantic triple ($Subject$, $Predicate$, $Object$) or ($Entity$, $Relationship$, $Entity$), where the entity could be any physical property or thing. For example, $\{SirajMunir, Study\_At, UniversityofUrbino\}$. A dynamic KG is a timestamped version of a static KG, denoted as $\{KG_1, KG_2, \ldots.., KG_N\}$. In our case, each relationship has an associated timestamp, such as $\{Author_A, met(timestamp), Author_B\}$.

B. COMPLEX NETWORK ANALYSIS

Complex Network Analysis (NA) allows for characterizing a given graph through a collection of metrics that provide insights from various statistical perspectives. Examples include degree, connectivity, centrality, modularity, etc. Dynamic Network Analysis is an extension of NA, with the notion of dynamicity being based on time. In other words, the evolution of a network is tied to a timeline. In this work, instead of dealing with simpler graphs, we model the dynamic KG as $\{KG_1, KG_2, KG_3, \ldots KG_N\}$, as mentioned in the previous sub-section. For detailed insights into static and dynamic network analysis, refer to the methodology section.

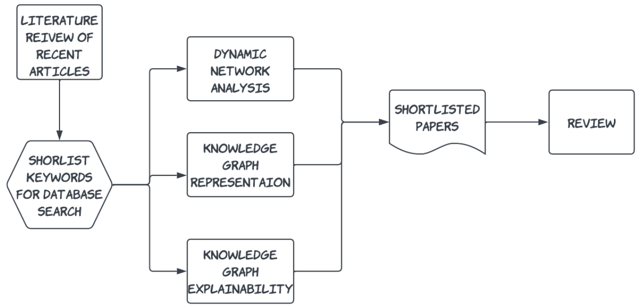

FIGURE 1. Literature review pipeline.

C. STATE OF THE ART

In this section, we will spotlight the latest contributions that align with our presented work. To shortlist articles, we applied the query criteria depicted in Figure 1. In selecting research articles, we exclusively relied on top-notch articles from reputable databases such as IEEE, ScienceDirect, Springer, etc. Given the plethora of articles, we considered only those published in the last 5 years.

As part of our filtration criteria, we selected three keywords: Knowledge Graph Representation, Knowledge Graph Explainability, and Dynamic Network Analysis. Additionally, we incorporated keyword extensions, including temporal KG analysis, spatiotemporal KG analysis, factual and counterfactual reasoning for KG, etc. Following the application of these criteria, we conducted a meticulous review, summarizing key points from the selected articles.

1) KNOWLEDGE GRAPH REASONING AND REPRESENTATION

KG is a hot research area, and several public and commercial solutions have recently leveraged its semantic power fantastically. For instance, KG completion and reasoning via link prediction, neighborhood prediction, and community detection have been explored [6], [16], [17]. To provide a snapshot of the current state of the art, we highlight relevant work in this section, focusing on KG completion, modeling, explainability, and representation.

2) KNOWLEDGE GRAPH REPRESENTATION AND REASONING (SYMBOLIC)

In a study by [18], a KG completion and reasoning approach for knowledge enrichment tasks was presented. Symbolic reasoning poses challenges, particularly in terms of scalability and interpretability. This study proposed an approach employing knowledge extraction, relational reasoning, and inconsistency checks. A comparative analysis validated the approach, outperforming formal methods. Another survey by [3] explored challenges in symbolic and neural symbolic reasoning, presenting a hybrid approach. Reference [5] proposed a deep learning model for symbolic representation and explainability, validated for cultural heritage use cases. A survey by [16] discussed present and future perspectives for symbolic KG reasoning, including related technologies. Reference [19] introduced a similar approach to the presented study. The proposed approach prunes textual representation and extracts rules from modeling KG.

3) KNOWLEDGE GRAPH AND STATIC GRAPH EXPLAINABILITY

In [20], a bi-kernel homophily-heterophily modeling approach for Graph Neural Networks (GNN) was presented, emphasizing its efficiency for consistent results. Reference [21] introduced a causal inference theory-driven approach for factual and counterfactual reasoning over GNN, focusing on factual reasoning and KG interpretation. Reference [22] proposed an interpretability-based attribution approach for GNN recommendations, and [2] presented a visual evaluation approach named GraphXAI, extending GNN explainability. In [23], a survey on link prediction and related tasks explored potential applications and the critical role of link prediction. Reference [24] identified research frontiers and hotspots for utilizing KG for recommendations. Reference [25] presented a KG-based reasoning question-answering system, achieving state-of-the-art results. Reference [26] introduced a KG clustering methodology based on information fusion, and [27] proposed a KG toolkit for data science applications. Reference [28] conducted a detailed survey of KG representation, highlighting hotspots, contributing countries, and top literature. In [29], a KG modeling framework was presented, utilizing information fusion for system modeling. Reference [30] introduced a unique methodology for information representation in KG, emphasizing a semi-autonomous approach. Another study [31] presented an interesting approach for interpretable and explainable deep learning models. The author emphasized how KG can add more value to decision-making and make black-box models translucent. Reference [32] work that is currently available as a pre-press version introduced a very interesting approach for interpreting KG. The authors presented a detailed framework comprised of three parts (i) mechanization layer (ii) composition layer, and (iii) assistance layer. Reference [33] presented an encouraging KG modeling approach for predicting and interpreting disease based on gene interaction. The presented approach integrates heterogeneous data sources for modeling KG. In contrast to the mentioned works the presented study differs as it integrates complex network analysis along with KG for interpretability and explainability.

4) KNOWLEDGE GRAPH COMPLETION

Reference [16] introduced a temporal KG completion approach incorporating link prediction, showing effectiveness for spatial-temporal link prediction tasks. Reference [34] introduced a box embedding approach for temporal KGs, intensely validating it for inference and expressivity. Reference [35] proposed the T-GAP model for path-based inference in static KG, outperforming baselines. Reference [36] proposed scoping for KG to enhance temporal representation, validated for various tasks. Reference [37] presented a 4-order tuple representational modeling for temporal KG, achieving superior performance. Reference [38] introduced the RLogic model for KG completion, utilizing deductive reasoning for recursive rule patterns.

5) DYNAMIC NETWORK ANALYSIS AND REPRESENTATION

In this section, we explore the current state-of-the-art for dynamic network representation and analysis, considering dynamic networks as presented in [7] and [39]. Recent works addressing challenges in dynamic network analysis have drawn inspiration from both machine and deep learning [40], [41], [42], [43], [44], [45], [46].

For example, in [40], dynamic metrics for assessing dynamic networks were introduced using a deep neural architecture to analyze temporal growth. The approach was validated with link prediction over real-world datasets. Reference [41] explored dynamic network connectivity patterns, providing insights into graph connectivity and statistical features. Reference [42] focused on neuroimaging and temporal community analysis, presenting a toolbox for MATLAB based on time-varying structural features. Reference [44] conducted a thorough survey on dynamic networks, covering aspects like taxonomy and definitions. In [45], a survey on data modeling and embedding strategies for dynamic networks was presented. Reference [46] introduced an approach for Graph Neural Networks (GNN) utilizing LSTM neural network architecture for temporal features, achieving successful link prediction. The literature reports significant results for dynamic networks across various analytical contexts. However, further advancements are anticipated, with researchers actively exploring this domain. The subsequent subsection will shed light on the methodological details.

III. METHODOLOGY

In this section, we delve into the methodological details of the presented workflow, which integrates our previous works [7], [30], [47]. The framework proposed in [30] focused on symbolic representation using First Order Logic (FOL) and Horn clause-driven rule extraction from raw data, presented as a Knowledge Graph (KG) for querying. While [7] addressed challenges of dynamic networks represented as KGs.

Our latest work introduced a federated scheme for querying over distributed decentralized KGs [47]. The intention of discussing the previous work and the inclusion of them in the present framework is to show the contributions that led us to this idea of extending KG's factual explainability. However, this work particularly focuses on the KG analysis phase. Figure 2 presents the end-to-end workflow, for better understandability we divided the framework into two parts: (i) information processing and (ii) information representation.

Phase 1: Information Processing

This phase is divided into three parts, serving to process raw data, extract symbolic representation, and prepare a KG for the subsequent phase.

1) **Data Acquisition:** Collection of data from real-world environments using sensors or camera feeds. In this study, as a use case, we utilized data collected from proximity sensors. Details of the dataset are mentioned in the subsequent subsection i.e., information processing.
2) **Information Fusion:** For the integration of collected data into a universal representation, that is transferable to symbolic rule extraction we aggregated acquired data from the previous phase.
3) **Symbolic Knowledge (Rule) Extraction:** This part utilizes the universal view prepared at the information fusion phase to extract first-order logic rules from a machine learning-based classification algorithm to establish decision boundaries. For more details on rule extraction please refer to our previous work [30].

These three phases complement the first part of the workflow. The details of phase 1 operations are elaborated in algorithm 1.

Algorithm 1 Algorithm for Phase 1

Input: Raw Data from the environment
Output: First-Order Logic Rules for KG Modeling
1: Import all raw data from n sources and transform them into a dataframe.
2: Aggregate data sources from prior step $\{d_1, d_{,2}, d_3, \dots .d_n\}$ into a universal representation D
3: Identify FOL rules using machine learning algorithm (CART, ITER, GridEx, REAL, Trepan) [30], [48]
4: Identify decision boundaries using rules and model KG involving human-in-loop.
5: **return** FOL Rules

Phase 2: Information Representation

Similarly, the second part (information representation) is divided into three parts: KG Federation, KG analysis, and federated querying, intending to aid KG representation analysis using Network Analysis (NA) metrics.

1) **KG Federation:** Distribution of information represented in KGs.
2) **KG Analysis:** Extraction of explainable details such as community structures, bridges, and ties. This work extends this phase by introducing homophily and heterophily analysis to aid factual reasoning of the given network represented as KG.
3) **Federated Querying:** Aggregation of knowledge stored in different chunks, making it queryable for the end-user. Note that this work does not emphasize the querying aspect.

The details of phase 2 operations are elaborated in algorithm 2.

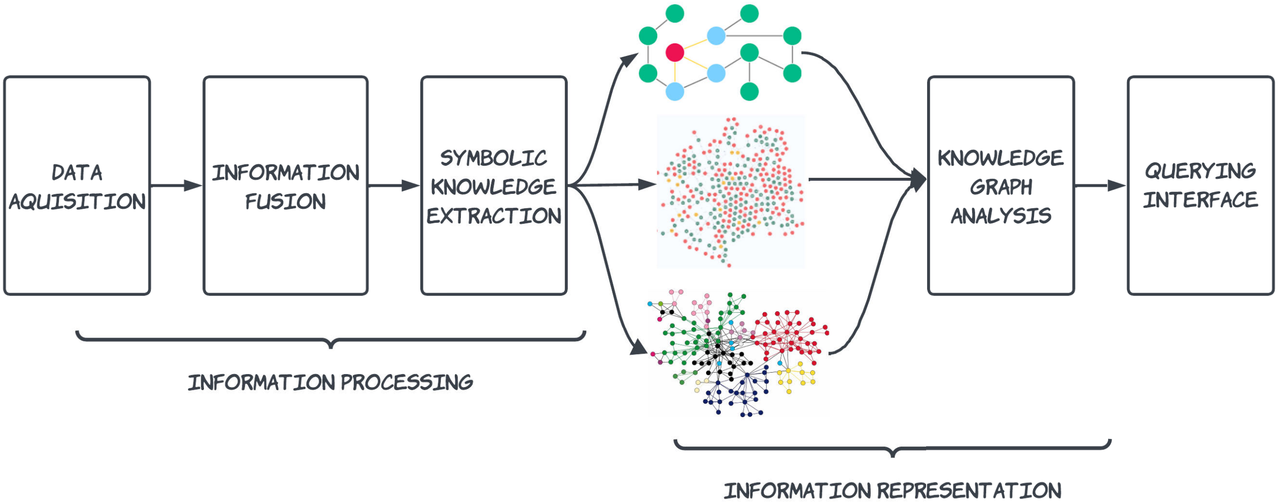

FIGURE 2. Workflow for knowledge graph management and analysis.

Algorithm 2 Algorithm for Phase 2

Input: Rules from Phase 1
Output: Knowledge Graph and Analysis
1: Identify rules as a decision boundary for community representation and pass it to the KG modeler.
2: Transform data rows as semantic triples $\{Subject, Predicate, Object\}$
3: **if** Multiple KG exists **then**
4: Federate them into a suitable number of shards as suggested in [47]
5: **else**
Continue with centralized KG. representation.
6: **end if**
7: Pass Modeled KG to KG Analysis Phase. See Algorithm 3.
8: Query for desired data.
9: **return KG**

Table 1 provides a list of mathematical symbols. The preceding discussion offers a general overview of the proposed workflow, justifying its potential extension for different domain applications. Subsequent sections will delve into a detailed discussion of this workflow in contrast to the presented study.

TABLE 1. Table of mathematical symbols.

Symbol	Description
A	Adjacency Matrix
I	Identity Matrix
U, V	Random Nodes
$n-1$	Number of nodes reachable to U
N	Total number of nodes in a graph
σ,d	Distance or count of nodes having shortest path
ΔQ	Modularity
m	size of graph
k	Weighted Sum
s, t	Random Nodes
γ	Neighbor
h_0	Heterophily Hyperparameter
D	Dynamic Network Rewiring Score
S, c, d	State, Centroid, and Dissimilarity Measure
h	Homophily Hyperparameter
A, B	Nodes Groups
P_{AB}	Homophily Probability

A. INFORMATION PROCESSING (DATA ACQUISITION, INFORMATION FUSION, AND SYMBOLIC KNOWLEDGE EXTRACTION)

In data acquisition, we relied on an open-source dataset available at [49] and [50]. The dataset is based on face-to-face interaction of authors attending international conferences held in 2016 and 2017, respectively. Data was collected with the help of RFID (Radio Frequency Identifier) based sensors with a proximity of approximately 1.5m. Along with social interactions, the dataset also includes a temporal track with a frequency of 20 seconds. Using author's association, the dataset also includes some demographic features like {Age, Seniority level (Education), background, or discipline, and Country} or semantically {Author_A, met(timestamp), Author_B } or {(Author_A) Has_Country(A), Has_Education(Ph.D), Has_Attended (ICCSS17)}. For our analysis, we divided the dataset into two parts i.e., social and temporal. First, we aggregated and transformed the tabular data into a symbolic representation using the scheme presented in [30]. The symbolic representation takes care of the information processing phase and generates rules based on First Order Logic (FOL) and Horn clauses. Then, based on the extracted rules and human-in-loop as fact-checkers and rule modelers, we modeled the KG.

```
'Country'(ID, `Level of Seniority`, `Level of Seniority.1`, `Level of Seniority.2`, `Level of Seniority.3`, Meeting, Met, Sex,
Timestamp, P1, P2, 1.0) :-
    `Level of Seniority` =< 5.5, `Level of Seniority.3` > 3.5.
'Country'(ID, `Level of Seniority`, `Level of Seniority.1`, `Level of Seniority.2`, `Level of Seniority.3`, Meeting, Met, Sex,
Timestamp, P1, P2, 3.22) :-
    `Level of Seniority` =< 5.5, `Level of Seniority.3` =< 3.5, `Level of Seniority` > 4.5.
'Country'(ID, `Level of Seniority`, `Level of Seniority.1`, `Level of Seniority.2`, `Level of Seniority.3`, Meeting, Met, Sex,
Timestamp, P1, P2, 3.0) :-
    `Level of Seniority` =< 5.5, `Level of Seniority.3` =< 3.5, `Level of Seniority` =< 4.5, `Level of Seniority` =< 3.5.
'Country'(ID, `Level of Seniority`, `Level of Seniority.1`, `Level of Seniority.2`, `Level of Seniority.3`, Meeting, Met, Sex,
Timestamp, P1, P2, 1.7) :-
    `Level of Seniority` =< 5.5, `Level of Seniority.3` =< 3.5, `Level of Seniority` =< 4.5, `Level of Seniority` > 3.5.
'Country'(ID, `Level of Seniority`, `Level of Seniority.1`, `Level of Seniority.2`, `Level of Seniority.3`, Meeting, Met, Sex,
Timestamp, P1, P2, 4.52) :-
    ID =< 241.5.
'Country'(ID, `Level of Seniority`, `Level of Seniority.1`, `Level of Seniority.2`, `Level of Seniority.3`, Meeting, Met, Sex,

'Output Label'(ID, `Level of Seniority`, `Level of Seniority.1`, `Level of Seniority.2`, `Level of Seniority.3`, Meeting, Met,
Sex, Timestamp, P1, P2, 1.0) :-
    ID in [5.99, 105.00].
'Output Label'(ID, `Level of Seniority`, `Level of Seniority.1`, `Level of Seniority.2`, `Level of Seniority.3`, Meeting, Met,
Sex, Timestamp, P1, P2, 3.0) :-
    ID in [5.99, 133.00].
'Output Label'(ID, `Level of Seniority`, `Level of Seniority.1`, `Level of Seniority.2`, `Level of Seniority.3`, Meeting, Met,
Sex, Timestamp, P1, P2, 3.0).
```

FIGURE 3. Symbolic rule representation.

TABLE 2. Summary of knowledge graph.

Nodes	464
Relationships	199,572. Label Types: Met, Has_Country, Has_Education, Has_Attended

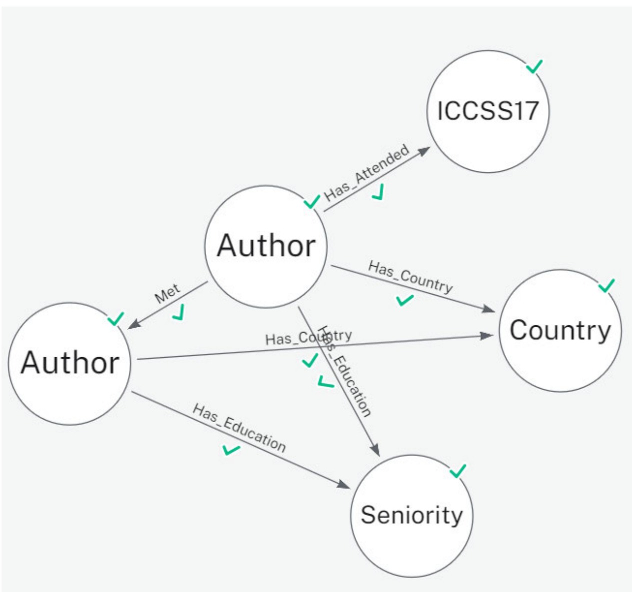

FIGURE 4. KG data model.

B. KNOWLEDGE GRAPH MODELING

For KG modeling, we utilized the rules extracted from the previous phase. The rules extracted from the symbolic knowledge extraction (rules) can be considered a decision boundary that helps identify the class of a given node in a graph. In other words, which parameter or what is measured makes up the particular class representation. The rules extracted from the symbolic representation are illustrated in Figure 3. As mentioned earlier these rules represent the classes. For instance, one of the sample rules states individuals with seniority levels 1, 2, and 3, and who have interacted once, possess IDs within the range of 6 to 105, and belong to a distinct class. Similarly, individuals who have interacted only once have seniority levels ranging from 3 to 5 belong to another distinct class, and so on.

Based on these rules, and human-in-loop as a domain expert we modeled the KG. For the modeling of KG, we adopted the data model presented in Figure 4. Notice that we named both conferences a single node label because based on the available dataset there was no clue to differentiate; so, for our analysis, we worked with the merged version. A brief statistical description of the obtained KG is presented in Table 2.

This helps us to federate the knowledge from social and temporal representation into a KG. Now with KG, we can also query and filter the interactions based on customized criteria(s). For example, we want to filter authors attending the conference held in a specific country. We can filter it by using the following query $\{Match(c : ICCSS17) \leftarrow [: Has\_Attended] - (a : Author) - [r : Has\_Country] -$ () return c,a,r}. If required, for upscaled analysis, we can also distribute data into a suitable number of shards and query them by the federated querying scheme proposed in [47].

C. KNOWLEDGE GRAPH ANALYSIS

In this section, we will discuss the fundamentals of KG analysis and how we contributed to factual explainability via reasoning. The foundational task of Network Analysis (NA) is to help end-users analyze the given network representation using statistical features. These approaches come in very handy because they help filter out interesting measures instead of just numbers. In this work, we utilized NA for static and dynamic analysis. The notion of dynamic NA is

FIGURE 5. Modularity-based community detection.

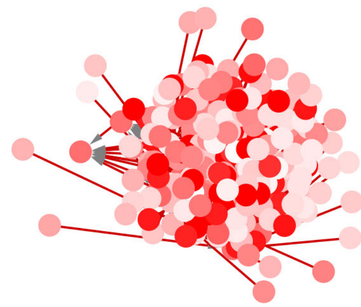

FIGURE 6. Dynamic node analysis based on interaction.

adopted from our previous work [7]. As per our study, simply relying on NA is not sufficient because statistics without semantics are incomplete. In other words, it is like you have identified an issue but you don't know how to track it down. Using semantic representation like KG helps you interlink the domain constraints via relationships. For ease of understanding, we have divided this section into the following sub-sections: (i) KG analysis as Static NA, (ii) KG analysis as Dynamic NA, and (iii) Factual reasoning of KG (homophily and heterophily).

1) STATIC NETWORK ANALYSIS

In this subsection, we will shed light on the details of NA. For in-depth analysis, we divided this phase into two parts: static and dynamic. For static NA, we simply divided our KG into two partitions: social interaction and temporal interaction. To identify communities of interest, we utilized the mathematical formulation presented in Equation 5. However, evaluating these communities requires a matrix representation along with closeness and betweenness centrality. To transform the network into matrix form and facilitate calculations, we applied the mathematical expressions presented in Equations 1, 2, and 3. Figure 5 presents the temporal and social interaction community partitions based on the modularity-based Louvain community detection algorithm [51]. Specifically, Figure 5A represents the temporal communities while Figure 5B represents the social interaction-based communities. Based on NA of social and temporal networks obtained from dataset [49], we identified 5 and 3 communities, respectively. As can be observed from the figures, the communities are overlapped and hence require further investigation. For further investigation of these communities and their interactivity, we applied the metrics of dynamic NA so that we can identify the most participating individuals based on their activity.

Equations 1-5 exhibit the mathematical formulation for evaluating modularity-based community detection presented in Figure 5.

$$Node\_Connection\_Matrix = \mathrm{A} + \mathrm{I} \tag{1}$$

where A and I are the adjacency and identity matrix [52].

$$Closeness\_Centrality(U) = \frac{n-1}{N-1}\frac{n-1}{\sum_{V=1}^{n-1} d(u,v)} \tag{2}$$

where d is a distance and $d(u, v)$ represents the shortest path between u and v, and $n-1$ is the number of nodes reachable to U and N is the total number of nodes in the graph [53].

$$Betwenness\_Centrality(V) = \sum_{s,t\in V} \frac{\sigma(s,t|V)}{\sigma(s,t)} \tag{3}$$

where V is the set of nodes, (s, t) and $\sigma(s, t|v)$ are the number of shortest paths [54].

$$Modularity = \Delta Q = \frac{k_{i,in}}{2m} - \gamma \frac{\Sigma_{tot} \cdot k_i}{2m^2} \tag{4}$$

Here, m represents the size of the graph, $k_{i,in}$ denotes the sum of weights from node i to the community, γ stands for the resolution parameter, k_i signifies the sum of weights of links incident to node i, and Σ_{tot} represents the sum of weights of links in the community [51].

2) DYNAMIC NETWORK ANALYSIS

For dynamic NA, we utilized the aggregated static view from static NA. Here, by static NA, we mean that for the initial analysis, we divided the network into two sub-graphs. These subgraphs were based on (i) temporal interaction and (ii) face-to-face or social interaction. As mentioned earlier, for dynamic NA, we adopted the notion that the dynamicity of the network depends on the interaction concerning time [55] and mathematically represented by:

$$D\ (Node\ X) = \frac{\sum_{n=1}^{S} d(x_i, c)}{S-1} \tag{5}$$

where D is a rewiring score used to measure the dynamic neighborhood of node $Node\ X$ in a dynamic network. A dynamic network is derived by a 3-dimensional matrix i.e., the extension of an adjacency matrix along with a network state or state space S. State space or network state is represented by the adjacencies view of the network snapshot i.e., $\{S_1, S_2, S_3, \ldots, S\}$. c is the relative mean centroid based on the variance within S, while d is the dissimilarity measure among nodes based on simple Euclidean distance. All of these measures are used together to calculate the dynamicity of the network [55].

We compared the network timestamps based on node activation and edge (interaction or relationship) activation. The results of the analysis are shown in Figures 6 and 7.

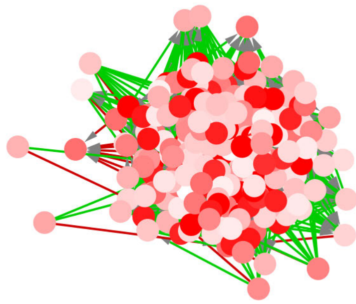

FIGURE 7. Dynamic edges analysis based on interaction.

Figure 6 shows the changes or dynamicity in the network concerning node activity, i.e., author or person mobility, and Figure 7 shows the dynamicity concerning edge or relationship. In Figures 6 and 7, the different color shades represent the dynamicity or activity of individual nodes and relationships. Darker shades represent more activity within the network.

3) FACTUAL EXPLAINABILITY

This sub-section highlights the utilization of the NA as a tool for confirming factual reasoning specifically homophily and heterophily. Homophily, as previously mentioned, is a measure confirming the characteristic likeness of correlated items. These items could be associated with any individual or physical entity (e.g., machine, sensor-based device, Internet of Things). On the contrary, heterophily is not generally represented by likeness.

In this study, we aim to confirm a factual interpretation or explanation of why any entity, in this case, a person (author), is a part of a certain community. In other words, like-minded authors tend to associate with other like-minded authors or heterophily, authors could be part of another community irrespective of their interests, which is mostly false in social networks. To prove this notion, we employed both homophily and heterophily as measures to observe the selected KG as a network based on four characteristics: age, seniority, country, and background. This type of homophilic analysis is also known as status-based homophily.

For robustness, we compared our network with two stochastic network analysis models: (i) Erdős-Rényi model and (ii) Block Model [49], [50]. Both models use a contact matrix for homophily analysis and an inverse contact matrix for heterophily analysis. Based on the earlier criteria, we divided all nodes into two groups A and B. The homophily measure can be represented by the symbol h and heterophily by h_0. Whereas h_{AA} represents the probability between group A members and h_{BB} represents the probabilities of group B members [50]. Whereas h_{AB} and h_{BA} are the complementary probabilities i.e., $h_{AB} = 1 - h_{AA}$ $h_{BA} = 1 - h_{BB}$. Furthermore, for feature selection, we assumed simple intuition i.e., homophily and heterophily are symmetry and complementary and can be represented by ($h = h_{AA} = h_{BB}$), ($h_{AB} = h_{BA} = 1 - h$) [50]. However, the value of

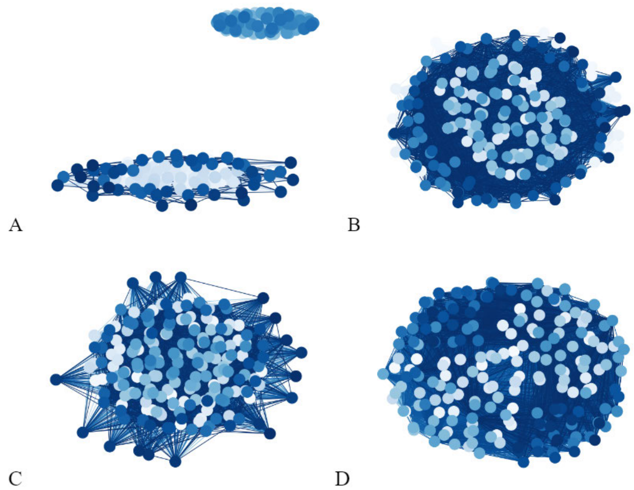

FIGURE 8. Homophily-dominated clustering of community features.

h ranges from 0 to 1, where $0 \geq$ h ≤ 0.5 means the nodes from one group are connected with a few specific groups (i.e., heterophily domination). If the value of $0.5 \geq h \leq 1$, this means the group of nodes is part of the same community and well-connected (homophily domination). If the value of $h = 0.5$, then the network is neutral or mixed. For the calculation of group-level homophily probability, we adopted a 4×4 matrix representing each group like P_{00} refers to the homophilic probability between Age to Age feature and so on, which is calculated by the following equations

$$h_{A,B} = \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \\ P_{30} & P_{31} & P_{32} & P_{33} \end{bmatrix} \tag{6}$$

However, each entry in the contact matrix is a node-level probability and is calculated by the following equation:

$$\text{Homophily Probability}(P_{i,j}) = \frac{h_{AB}(i,j)K_i}{\sum_e h_{BA}(e,j)K_e} \tag{7}$$

where i and j are connecting nodes within a specific group e.g. Age or Country and (e) represents the total edge count and K_i, K_e are degree counts for all edges in i. Note that $\{\ \forall\ values\ of\ h_{A,B}(i,j) \equiv h\}$. Where h represents the homophily hyperparameter for node i connecting node j in groups A and B [50].

$$P_{AB} = \frac{(P_A P_B)^{1/2}}{(P_A P_B)^{1/2} + [(1 - P_A)(1 - P_B)]^{1/2}} \tag{8}$$

where P_{AB} is the mean homophilic probability between groups A and B [50].

For inverse Block Model calculation i.e., heterophily, we just took the inverse of Block Model calculations, i.e., $(1 - h_{A,B})$. For Erd˝s-Rényi model calculations, we utilized the same contact matrix used for the Block Model. The subsequent section will briefly explain the results of the proposed methodology. The details of the KG analysis are elaborated in algorithm 3.

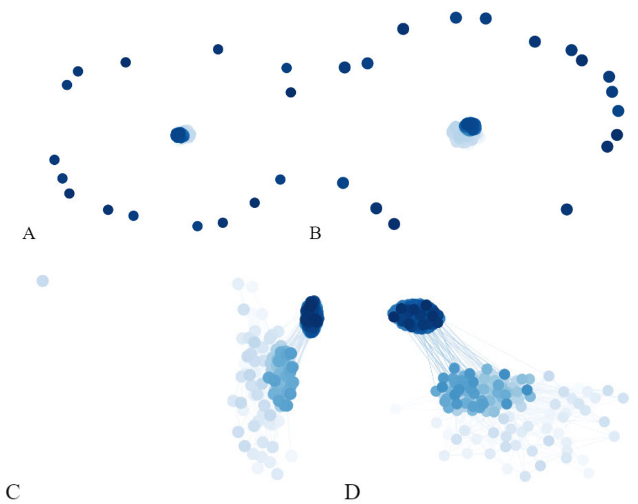

FIGURE 9. Heterophily-dominated clustering of community features.

Algorithm 3 Algorithm for KG Analysis

Input: Knowledge Graph Representation
Output: Factual Explanation and Interpretation

1: Identify communities of interest based on equation 4. To cross-check the fact-based communities extracted from algorithm 2.
2: Identify dynamic nodes and edges based on temporal activities as mentioned in equation 5.
3: Combine all temporal snapshots of KG and transform them into a universal KG for extraction of factual interpretations.
4: For homophily and heterophily analysis probability
   1) Divide KG into two node groups A and B. Where $(h = h_{AA} = h_{BB})$, $(h_{AB} = h_{BA} = 1 - h)$
   2) Evaluated the homophily probabilities of the chosen groups by 4×4 parametric matrix as mentioned in equations 6-8.
   3) Make 3 variants for analysis i.e., homophily-dominated, heterophily-dominated, and mixed.
5: Cross-validate the results by correlation analysis (Block model vs Erdős-Rényi model), Preferential Attachment, and Z-score.
6: **return** Knowledge Graph Analysis Results as community or groups.

D. EXPERIMENTAL SETUP

For the presented study, we utilized Python programming language (numpy, pandas, networkX, and matplotlib), Gephi, and Cytoscape (an open-source network analysis platform). For computing resources, we used Intel Core-i7 10^{th} Generation. As mentioned earlier for this study we used an open-source dataset [49].

IV. RESULTS AND DISCUSSION

In this section, we will discuss in detail the results we achieved from the methodology mentioned earlier. For validation, we tested the model with two variations: (i) homophily domination and (ii) heterophily domination. Homophily domination refers to clustering or grouping, i.e., a community of authors based on likeness, and heterophily domination refers to a sparse group of authors having weak communication and hence less likely to be in the same cluster or community.

In this work, we implemented a community-level Block Model. There are other variants [49], [56] that utilize edge-level or relationship-level Block Models. But for simplicity, we adhere to the community level. As mentioned earlier, we focused only on four features (Age, Seniority, Country, and Language), and for the calculation of heterophily, we considered the inverse homophily matrix.

Figure 8 shows the homophily-dominated clustering of each feature respectively where A, B, C, and D refer to Age, Seniority, Country, and Language.

Figure 9 shows the heterophily-dominated features A, B, C, and D clustering similar to Figure 8. However, it is observable that only country and language-based features have a small number of clusters. This clustering gives a glimpse to cross-validate and ensure analysis; we incorporated Block and Erdős-Rényi model. As mentioned earlier, the Block Model is based on a 4×4 parametric matrix model, and Erdős-Rényi is considered a null model. Hence, to further investigate the analysis from both models, we model a correlation matrix shown in Figure 10. Where Figures 10 A, 10 B, and 10 C show the correlations for each feature based on the null model (Erdős-Rényi) in contrast Figures 10 D, 10 E, and 10 F show the correlations based on the Block Model.

However, to identify true homophily and heterophily, we made three variants of the correlation matrix, i.e., homophily dominated, heterophily dominated, and neutral. The correlation analysis results show that the KG has a homophilic nature. To confirm and validate that the presented analysis is true, we also implemented the preferential attachment algorithm presented in [57]. The preferential attachment algorithm suggests that famous entities in the network will receive more links as the network grows. Figure 11 shows the preferential attachment in a score-wise manner. This highlights that due to high homophily, more than half of the authors will connect to more authors.

Hence, the link prediction analysis based on the preferential attachment algorithm [57] confirms the correlation analysis results, i.e., the given KG is homophilic. Further, as an additional layer of the test, we incorporated a z-score measure A.K.A standard score in statistics that tells us how a specific data point(s) are relational to the mean. Figures 12A, 12B, and 12C present the z-score measure as a histogram. It is interesting as it also highlights some heterophilic nature of communities which was also prominent in some correlation cases. Similar to the previous analysis we made three variants i.e., homophily dominated, heterophily dominated, and mixed for the z-score measure.

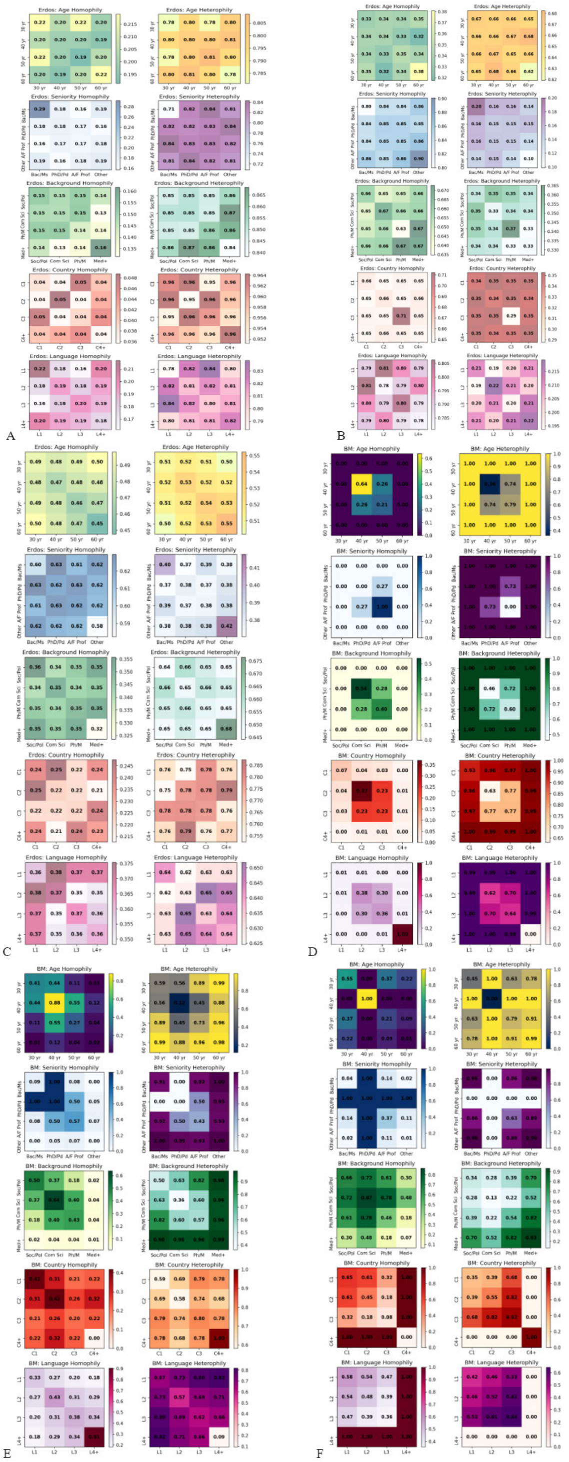

FIGURE 10. Correlation analysis of homophily and heterophily.

V. LIMITATIONS AND FUTURE WORK

This section addresses the limitations of our work and outlines potential directions for future exploration. The integrated approach proposed in this study combines symbolic rule extraction, Network Analysis (NA), and factual reasoning for Social Mobility Knowledge Graph (KG). We also discuss KG modeling and querying. To strengthen and validate our results, we conducted various experiments,

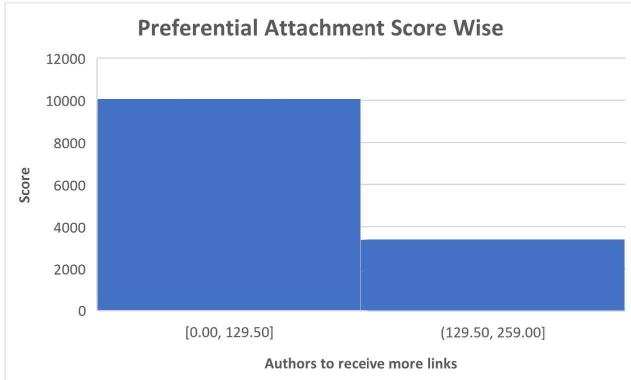

FIGURE 11. Preferential attachment score-wise analysis.

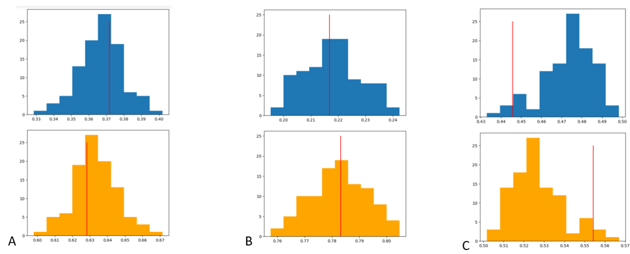

FIGURE 12. Z-score analysis.

including link prediction, correlation analysis, homophily, and heterophily analysis. This study is beneficial for behavioral analysis, hotspot identification, and profiling studies where we are interested in observing user(s) interaction(s) as a community. For example, digital advertisement, and marketing, workspace profiling, citizen profiling, and well-being, E-Government, etc.

We have conducted a data analysis, which confirms that the Social Mobility KG has a homophilic nature. However, we acknowledge that this observation may not be universally applicable. Therefore, it is important to explore multiple mobility KGs to corroborate our factual conclusions. Unlike some studies, we didn't introduce any novel matrices for factual reasoning. Instead, we relied on stochastic homophily and heterophily. Our conclusion is reasonable, but further investigation across different KG genres is required to assess stability (robustness) and correctness (truthfulness). Our study didn't employ any KG embedding scheme(s) for learning vector space. However, it is crucial to devise baseline model(s) that could be applied to integrated studies like the one presented.

Another avenue worth exploring is the application of Graph Neural Networks (GNNs) within our framework. Existing literature [2], [5], [6], [22], [26] has explored GNNs for explainability, but their application to KGs and utilization of homophily and heterophily-based metrics for reasoning remain unexplored areas. KG reasoning and embedding itself is a compelling research topic.

Furthermore, there is a need to explore NA and its metrics for reasoning in KGs, especially in temporal and spatiotemporal settings as discussed in [2] and [58]. This work presented a factual analysis, indicating that like-minded individuals tend to associate. In future work, we intend to improve the following (i) improvise supervised symbolic rule extraction by newer methods and without humans in the loop, (ii) Enhance the mathematical modeling of the block model for dealing with the complex feature set. Counterfactual reasoning is a less mature research topic, but it represents another intriguing direction, much like attention-based learning such as reinforcement or transfer learning.

Additionally, the integration of Large Language Models (LLMs) and Retrieval Augmented Generation (RAG) with real-world graph networks is an underexplored area. Investigating different integration schemes and relevant evaluation metrics is a crucial research avenue. Graph representation and network science are essential topics for achieving better explainability. The subsequent section will provide concluding remarks.

VI. CONCLUSION

This work introduced an integrated approach using complex network analysis as a tool for understanding factual reasoning. We combined symbolic representation with KG modeling to enhance the ground truth representation. Furthermore, we introduced homophily and heterophily analysis and link prediction to validate factual reasoning. Our framework enables the extraction of rules from raw data and the modeling of a KG, ready for network analysis-based measures for factual reasoning. However, as highlighted in our discussion of future work, some aspects require further attention and exploration.

Explainable Artificial Intelligence (XAI) is at the forefront of research and industry, revealing intricate details that may appear complex from a top-level perspective. The complexity hidden within data representation and algorithms necessitates new approaches to revealing these details. Consequently, further studies, akin to those presented in this article, are essential for advancing the understanding of these complex systems.

REFERENCES

[1] Q. V. Liao and K. R. Varshney, "Human-centered explainable AI (XAI): From algorithms to user experiences," 2021, *arXiv:2110.10790*.
[2] C. Agarwal, O. Queen, H. Lakkaraju, and M. Zitnik, "Evaluating explainability for graph neural networks," *Sci. Data*, vol. 10, no. 1, p. 144, Mar. 2023.
[3] S. Chang, E. Pierson, P. W. Koh, J. Gerardin, B. Redbird, D. Grusky, and J. Leskovec, "Mobility network models of COVID-19 explain inequities and inform reopening," *Nature*, vol. 589, no. 7840, pp. 82–87, Jan. 2021.
[4] N. Maghawry, S. Ghoniemy, E. Shaaban, and K. Emara, "An automatic generation of heterogeneous knowledge graph for global disease support: A demonstration of a cancer use case," *Big Data Cognit. Comput.*, vol. 7, no. 1, p. 21, Jan. 2023.
[5] N. Díaz-Rodríguez, A. Lamas, J. Sanchez, G. Franchi, I. Donadello, S. Tabik, D. Filliat, P. Cruz, R. Montes, and F. Herrera, "EXplainable neural-symbolic learning (X-NeSyL) methodology to fuse deep learning representations with expert knowledge graphs: The MonuMAI cultural heritage use case," *Inf. Fusion*, vol. 79, pp. 58–83, Mar. 2022.
[6] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, Jan. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666651021000012
[7] S. Munir, S. Ferretti, and R. A. Shams, "A framework for knowledge representation integrated with dynamic network analysis," in *Proc. Int. Conf. Emerg. Trends Appl. Artif. Intell.* (Lecture Notes in Networks and Systems), Springer.
[8] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociology*, vol. 27, no. 1, pp. 415–444, Aug. 2001, doi: 10.1146/annurev.soc.27.1.415.
[9] C. Lozares, J. M. Verd, I. Cruz, and O. Barranco, "Homophily and heterophily in personal networks. From mutual acquaintance to relationship intensity," *Qual. Quantity*, vol. 48, no. 5, pp. 2657–2670, Sep. 2014.
[10] M. Pósfai and A.-L. Barabási, *Network Science*. Cambridge, U.K.: Cambridge Univ. Press, 2016.
[11] M. Newman, *Networks*. London, U.K.: Oxford Univ. Press, 2018.
[12] A. S. D. Mata, "Complex networks: A mini-review," *Brazilian J. Phys.*, vol. 50, no. 5, pp. 658–672, Oct. 2020.
[13] D. Billington, *Factual and Plausible Reasoning*. Brisbane, QL, Australia: College Publications, 2019.
[14] A. Singhal, *Introducing the Knowledge Graph: Things, Not Strings*, May 2012. [Online]. Available: https://blog.google/products/search/introducing-knowledge-graph-things-not/
[15] M. E. J. Newman, "Modularity and community structure in networks," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 23, pp. 8577–8582, Jun. 2006, doi: 10.1073/pnas.0601602103.
[16] J. Zhang, S. Liang, Z. Deng, and J. Shao, *Spatial–Temporal Attention Network for Temporal Knowledge Graph Completion*. Taipei, Taiwan: Springer, 2021.
[17] M. Molinaro and G. Orzes, "From forest to finished products: The contribution of Industry 4.0 technologies to the wood sector," *Comput. Ind.*, vol. 138, Jun. 2022, Art. no. 103637.
[18] H. Chen, S. Deng, W. Zhang, Z. Xu, J. Li, and E. Kharlamov, "Neural symbolic reasoning with knowledge graphs: Knowledge extraction, relational reasoning, and inconsistency checking," *Fundam. Res.*, vol. 1, no. 5, pp. 565–573, Sep. 2021.
[19] Y. Pan, J. Liu, L. Zhang, and Y. Huang, "Incorporating logic rules with textual representations for interpretable knowledge graph reasoning," *Knowl.-Based Syst.*, vol. 277, Oct. 2023, Art. no. 110787. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705123005373
[20] L. Du, X. Shi, Q. Fu, X. Ma, H. Liu, S. Han, and D. Zhang, "GBK-GNN: Gated bi-kernel graph neural networks for modeling both homophily and heterophily," in *Proc. ACM Web Conf.* New York, NY, USA: Association for Computing Machinery, Apr. 2022, pp. 1550–1558, doi: 10.1145/3485447.3512201.
[21] J. Tan, S. Geng, Z. Fu, Y. Ge, S. Xu, Y. Li, and Y. Zhang, "Learning and evaluating graph neural network explanations based on counterfactual and factual reasoning," in *Proc. ACM Web Conf.* New York, NY, USA: Association for Computing Machinery, 2022, pp. 1018–1027.
[22] B. Sanchez-Lengeling, J. Wei, B. Lee, E. Reif, P. Y. Wang, W. W. Qian, K. McCloskey, L. Colwell, and A. Wiltschko, "Evaluating attribution for graph neural networks," in *Proc. 34th Int. Conf. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2020, pp. 5898–5910.
[23] A. Rossi, D. Barbosa, D. Firmani, A. Matinata, and P. Merialdo, "Knowledge graph embedding for link prediction: A comparative analysis," *ACM Trans. Knowl. Discovery Data*, vol. 15, no. 2, pp. 1–49, Jan. 2021, doi: 10.1145/3424672.
[24] B. Shao, X. Li, and G. Bian, "A survey of research hotspots and frontier trends of recommendation systems from the perspective of knowledge graph," *Expert Syst. Appl.*, vol. 165, Mar. 2021, Art. no. 113764. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417420305881
[25] M. Yasunaga, H. Ren, A. Bosselut, P. Liang, and J. Leskovec, "QA-GNN: Reasoning with language models and knowledge graphs for question answering," 2021, *arXiv:2104.06378*.
[26] J. Tang, Y. Liu, K.-Y. Lin, and L. Li, "Process bottlenecks identification and its root cause analysis using fusion-based clustering and knowledge graph," *Adv. Eng. Informat.*, vol. 55, Jan. 2023, Art. no. 101862. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1474034622003202
[27] F. Ilievski, D. Garijo, H. Chalupsky, N. T. Divvala, Y. Yao, C. Rogers, R. Li, J. Liu, A. Singh, D. Schwabe, and P. Szekely, "KGTK: A toolkit for large knowledge graph manipulation and analysis," 2020, *arXiv:2006.00088*.
[28] X. Chen, H. Xie, Z. Li, and G. Cheng, "Topic analysis and development in knowledge graph research: A bibliometric review on three decades," *Neurocomputing*, vol. 461, pp. 497–515, Oct. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231221009528

[29] J. Lü, G. Wen, R. Lu, Y. Wang, and S. Zhang, "Networked knowledge and complex networks: An engineering view," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 8, pp. 1366–1383, Aug. 2022.
[30] S. Munir and S. Ferretti, "Towards symbolic representation-based modeling of temporal knowledge graphs," in *Proc. Int. Conf. Smart Appl., Commun. Netw. (SmartNets)*, Jul. 2023, pp. 1–8.
[31] M. Gaur, K. Faldu, and A. Sheth, "Semantics of the black-box: Can knowledge graphs help make deep learning systems more interpretable and explainable?" *IEEE Internet Comput.*, vol. 25, no. 1, pp. 51–59, Jan. 2021.
[32] Y. Chudasama, D. Purohit, P. D. Rohde, J. Gercke, and M.-E. Vidal, "InterpretME: A tool for interpretations of machine learning models over knowledge graphs," *Semantic Web*, pp. 1–21, Jan. 2024.
[33] A. Renaux, C. Terwagne, M. Cochez, I. Tiddi, A. Nowé, and T. Lenaerts, "A knowledge graph approach to predict and interpret disease-causing gene interactions," *BMC Bioinf.*, vol. 24, no. 1, p. 324, Aug. 2023.
[34] J. Messner, R. Abboud, and I. I. Ceylan, "Temporal knowledge graph completion using box embeddings," in *Proc. AAAI Conf. Artif. Intell.*, Jun. 2022, vol. 36, no. 7, pp. 7779–7787.
[35] J. Jung, J. Jung, and U. Kang, "Learning to walk across time for interpretable temporal knowledge graph completion," in *Proc. 27th ACM SIGKDD Conf. Knowl. Discovery Data Mining*. New York, NY, USA: ACM, Aug. 2021, pp. 786–795.
[36] J. Leblay, M. W. Chekol, and X. Liu, "Towards temporal knowledge graph embeddings with arbitrary time precision," in *Proc. 29th ACM Int. Conf. Inf. Knowl. Manage.*, 2020, pp. 685–694.
[37] P. Shao, D. Zhang, G. Yang, J. Tao, F. Che, and T. Liu, "Tucker decomposition-based temporal knowledge graph completion," *Knowl.-Based Syst.*, vol. 238, Feb. 2022, Art. no. 107841.
[38] K. Cheng, J. Liu, W. Wang, and Y. Sun, "RLogic: Recursive logical rule learning from knowledge graphs," in *Proc. 28th ACM SIGKDD Conf. Knowl. Discovery Data Mining*. New York, NY, USA: ACM, Aug. 2022, pp. 179–189.
[39] F. Harary and G. Gupta, "Dynamic graph models," *Math. Comput. Model.*, vol. 25, no. 7, pp. 79–87, Apr. 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0895717797000502
[40] P. Goyal, S. R. Chhetri, and A. Canedo, "dyngraph2vec: Capturing network dynamics using dynamic graph representation learning," *Knowl.-Based Syst.*, vol. 187, Jan. 2020, Art. no. 104816, doi: 10.1016/j.knosys.2019.06.024.
[41] A. N. Khambhati, A. E. Sizemore, R. F. Betzel, and D. S. Bassett, "Modeling and interpreting mesoscale network dynamics," *NeuroImage*, vol. 180, pp. 337–349, Oct. 2018.
[42] A. E. Sizemore and D. S. Bassett, "Dynamic graph metrics: Tutorial, toolbox, and tale," *NeuroImage*, vol. 180, pp. 417–427, Oct. 2018.
[43] J. Skarding, B. Gabrys, and K. Musial, "Foundations and modeling of dynamic networks using dynamic graph neural networks: A survey," *IEEE Access*, vol. 9, pp. 79143–79168, 2021.
[44] Y. Xie, C. Li, B. Yu, C. Zhang, and Z. Tang, "A survey on dynamic network embedding," 2020, *arXiv:2006.08093*.
[45] R. Triguero-Ocaña, B. Martínez-López, J. Vicente, J. A. Barasona, J. Martínez-Guijosa, and P. Acevedo, "Dynamic network of interactions in the wildlife-livestock interface in Mediterranean Spain: An epidemiological point of view," *Pathogens*, vol. 9, no. 2, p. 120, Feb. 2020.
[46] J. Chen, X. Wang, and X. Xu, "GC-LSTM: Graph convolution embedded LSTM for dynamic network link prediction," *Appl. Intell.*, vol. 52, no. 7, pp. 7513–7528, May 2022.
[47] S. F. S. Munir, "Towards federated decentralized querying with KG," in *Proc. Int. Conf. Comput. Sci. Comput. Intell.*, Dec. 2023.
[48] F. Sabbatini, G. Ciatto, R. Calegari, and A. Omicini, "On the design of PSyKE: A platform for symbolic knowledge extraction," in *Proc. WOA*, 2021, pp. 29–48.
[49] M. Génois, M. Zens, C. Lechner, B. Rammstedt, and M. Strohmaier, "Building connections: How scientists meet each other during a conference," 2019, *arXiv:1901.01182*.
[50] F. Karimi, M. Génois, C. Wagner, P. Singer, and M. Strohmaier, "Homophily influences ranking of minorities in social networks," *Sci. Rep.*, vol. 8, no. 1, p. 11077, Jul. 2018.
[51] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech., Theory Exp.*, vol. 2008, no. 10, Oct. 2008, Art. no. P10008, doi: 10.1088/1742-5468/2008/10/p10008.
[52] P. Bonacich, "Power and centrality: A family of measures," *Amer. J. Sociol.*, vol. 92, no. 5, pp. 1170–1182, Mar. 1987.
[53] U. Brandes, "On variants of shortest-path betweenness centrality and their generic computation," *Social Netw.*, vol. 30, no. 2, pp. 136–145, May 2008.
[54] W. Stanley and F. Katherine, *Social Network Analysis: Methods and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1994.
[55] I. H. Goenawan, K. Bryan, and D. J. Lynn, "DyNet: Visualization and analysis of dynamic molecular interaction networks," *Bioinformatics*, vol. 32, no. 17, pp. 2713–2715, Sep. 2016.
[56] L. Duvivier, R. Cazabet, and C. Robardet, "Edge based stochastic block model statistical inference," in *Complex Networks & Their Applications IX*, R. M. Benito, C. Cherifi, H. Cherifi, E. Moro, L. M. Rocha, and M. Sales-Pardo, Eds. Cham, Switzerland: Springer, 1007, pp. 462–473.
[57] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, Jan. 2003, doi: 10.1137/s003614450342480.
[58] S. Munir, R. A. S. Malick, S. I. Jami, G. Ahmed, S. Khan, and J. J. P. C. Rodrigues, "An integrated approach: Using knowledge graph and network analysis for harnessing digital advertisement," *Multimedia Tools Appl.*, vol. 82, no. 6, pp. 8883–8898, Mar. 2023.

SIRAJ MUNIR received the B.S. degree in computer sciences from Indus University, in 2016, and the M.S. degree in computer sciences from Muhammad Ali Jinnah University, Karachi, Pakistan, in 2019. He is currently pursuing the Ph.D. degree in computer science with the University of Urbino Carlo Bo, Urbino, Italy. He has published in well-reputed journals and conferences. He is exploring reasoning or explainability-driven graph neural networks, knowledge graphs, and network analysis in spatio-temporal settings. His research interests include knowledge graphs, information representation and modeling, complex and social network analysis, machine learning, and deep learning.

RAUF AHMED SHAMS MALICK received the Ph.D. degree from the University of Karachi. He has been a Visiting Scholar with the National Institute of Genetics (NIG), Japan, and UCLA, USA. He is currently affiliated with the Department of Computer Science, National University of Computer and Emerging Sciences, as an Assistant Professor, continuing research in the specialized scientific area of computer science, complex networks, social computing, bioinformatics, and integrated systems. He has founded several companies with state-of-the-art products related to social media, location-based analytics, and organizational networks. He is involved in complex system research and pursuing problems in biological networks, networked economics, and personality traits. He has a distinguished background in designing novel solutions for complex systems. He has authored several articles along with chapters in different books.

STEFANO FERRETTI received the Laurea (summa cum laude) and Ph.D. degrees in computer science from the University of Bologna, in 2001 and 2005, respectively. He was an Associate Professor with the Department of Computer Science and Engineering, University of Bologna. Since 2020, he has been an Associate Professor with the Department of Pure and Applied Sciences, University of Urbino Carlo Bo. His current research interests include distributed systems, complex networks, data science, fintech, and blockchain technologies, multimedia communications, and hybrid and distributed simulation. He is on the editorial board of the *Simulation Modelling Practice and Theory* (Elsevier) and *Encyclopedia of Computer Graphics and Games* (Springer). He is also on the Technical Committee of *Computer Communications* (Elsevier) and *Online Social Networks and Media* (Elsevier). He acted as an editor of special issues in other international journals (i.e., Wiley CPE and Elsevier ComCom).