IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 20 May 2022, accepted 16 July 2022, date of publication 26 July 2022, date of current version 4 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3194063

SURVEY

Knowledge Graphs in Education and Employability: A Survey on Applications and Techniques

YOUSRA FETTACH[1], MOUNIR GHOGHO[1,2], AND BOUALEM BENATALLAH[3,4], (Fellow, IEEE)

[1]TIC Laboratory, College of Engineering & Architecture, International University of Rabat, Sale 11103, Morocco
[2]Faculty of Engineering, University of Leeds, Leeds LS2 9JT, U.K.
[3]School of Computing, Dublin City University, Dublin 9, D09 W6Y4 Ireland
[4]School of Computer Science and Engineering, UNSW Sydney, Sydney, NSW 2052, Australia

Corresponding author: Yousra Fettach (yousra.fettach@uir.ac.ma)

ABSTRACT Studies on the relationship between education and employability are of paramount importance for policy makers, training institutions, companies and students. The availability of large scale data on the Internet, such as online job ads, has been leveraged to better investigate this relationship. Recent studies have shown that representing this data using knowledge graphs can be useful in identifying the job market needs and establishing better assessment methods. Such studies also highlight the skill mismatch between education and the job market and help mitigate it. In this paper, we provide a comprehensive review of the applications of knowledge graphs in education and employability. We review knowledge graph based frameworks that have been proposed in previous work on different aspects of education and employability. The survey introduces new taxonomies for these applications and their methodologies, and provides a thorough outlook on promising research directions. We also present a use case to illustrate how knowledge graphs can be used to investigate the relationship between education and the job market.

INDEX TERMS Educational technologies, employability, knowledge graphs, representation learning.

I. INTRODUCTION

Education has been an important tool for economic and social growth throughout the modern era. Higher education credentials in particular are considered by many people to be essential to enter the labor market. However, recent research demonstrated that sometimes higher education falls short in fulfilling the needs of the job market. Therefore, the issue of skill disparities between the job market and education has received much attention from policy makers, higher education institutions, employers as well as researchers. With the availability of massive amounts of structured and unstructured data engendered by the digital transformation, online content such as job ads, massive open online courses and resumes, has been incorporated in order to examine and analyze the issue.

The increased availability of large amount of data of different types also impelled researchers to leverage semantic and comceptual representation. That led to the birth of knowledge graphs (KG). A KG is a structured representation of facts, consisting of entities, relationships, and semantic descriptions [1]. KGs are core component for several information systems which require access to structured knowledge [2]. With the support of KGs as a representation tool for education and labor market entities, one can unravel correlations across these entities and construct semantic connections between them in order to turn data into usable knowledge.

Previous work proposed numerous methods based on data analysis to improve education by learning student behavior [3] or building recommender systems for students [4] for a better academic performance and education-to-job

The associate editor coordinating the review of this manuscript and approving it for publication was Ali Shariq Imran.

matching [5]. Several studies investigated ways to evaluate the job market using job representations [6], presenting taxonomies to help job seekers [7] and predicting job market trends [8]. Furthermore, many large-scale KGs have been developed by various academic and corporate institutions [9], [10].

This review is based on the analysis of research papers on the use of knowledge graphs as a tool to improve education and employability for job seekers and employers. Based on this analysis, we suggest a unified KG-based framework for education and employability. The main contributions of this paper are as follows:

- A comprehensive survey of KG applications in education.
- A comprehensive survey of KG applications for employability.
- A unified KG-based framework for education and employability.
- An outlook on future directions.
- A case study for exploring the relationship between education and the job market.

A. RELATED SURVEYS

Previous survey papers on KGs mainly focus on technical aspects [1], KG applications in general [2] and their embeddings [11]. Surveys like [12] and [13] provide general overviews of research from the perspective of job market evaluation using data analysis methods. Other literature reviews highlighted the potential of KGs in education [14], [15]. Our survey combines all these aspects and dives deeper into the applications of KGs in education and the techniques behind them. Besides, our paper provides a comprehensive review of KG-based methodologies for employability in addition to showcasing the role of KGs in evaluating the link between education and the job market.

B. SURVEY ORGANIZATION

In Section II, we describe the background of KG methodologies for education and employability. Section III discusses the KG techniques in education and employability. KG applications in education and employability are discussed in Section IV and V, respectively. In Section VI, we elaborate on the importance of building a unified framework for representing education and employability data in a single knowledge graph. This framework is illustrated using a use case in Section VII. Future research directions are discussed in VIII and conclusions are drawn in Section IX.

II. BACKGROUND

Studies took different directions in order to use semantic metadata to improve education. One of these directions is exploring ways to automatically define prerequisite relationships between concepts. From a pedagogical point of view, a prerequisite is a dependency relation that states which concepts a student has to learn before moving to the next one [16]–[18]. Several approaches were utilized to tackle this problem. One of the first studies in this realm was to enable prerequisite prediction for within-university and cross-university settings. The initial intuition of the work was the fact that courses may be specific to one institution whereas concepts are shared across different providers [19]. Some of the approaches adopted supervised learning to perform binary classification, using a set of features extracted via natural language processing methods, in order to automatically identify prerequisites [20]. Pan *et al.* [17] have implemented different classification methods using novel contextual and structural features extracted from MOOCs. Authors of [16] applied active learning based on graph-based features for representing concept pairs. In [21], the authors claimed that the problem of determining an effective learning path from a corpus of documents depends on the accurate identification of concepts that are learning objectives and the identification of concepts that are prerequisites. The authors of [21] then implemented a supervised machine learning approach for annotation based on multiple extracted local features and contextual ones. Other studies focused on the relationship between learning resources instead of concepts [22]. In [23], the authors suggested an approach to acquire concepts maps from textbooks, where concepts are associated with prerequisite relationships.

Various studies were established in order to assist job seekers for the purpose of finding a good fit [24], [25]. Graphs were first incorporated into these methodologies with graph-based job recommender systems [26]. In [27], the authors created a job transition network that presented the talent flow between organizations, which was then used to build a talent circle detection model for talent exchange prediction. Other studies presented methodologies that aimed to create new taxonomies or enhance existing ones for job-skill matching [7].

Overall, KGs started getting more attention in recent years [28] and being incorporated in various tools for education and job market assessment. The next section represents a brief report on KG techniques and how they are applied in education and employability.

III. KG TECHNIQUES IN EDUCATION AND EMPLOYABILITY

Knowledge representation went through many iterations based on frame semantics, which is the idea that understanding one word relies on the understanding of the knowledge related to it, representing coherent structure of related concepts. The concretization of this concept in computational linguistics can be found in the Semantic Web and KGs. The Semantic Web started by developing the concepts of the technical standards such as Uniform Resource Identifier (URI), Resource Description Framework (RDF) and Web Ontology Language (OWL). The concept of Linked Data started as linking datasets to each other in the Semantic Web to make them as a unified large KG. In 2012, the concept of KGs became more popular as Google launched it and utilized the semantic knowledge in web search in order to identify and differentiate entities in addition to providing links to related entities for the purpose of creating more of a connected and uniform search results.

There are many technical aspects in KGs [1]. In this section we try to tackle the most prominent ones in the field of education and employability.

A. KNOWLEDGE ACQUISITION

In KGs, knowledge acquisition aims to extract knowledge from unstructured text or various other structured or semi structured sources. It also aims at completing established KGs and recognizing entities and relations. Knowledge acquisition consists of many tasks including entity recognition and alignment and relation extraction [1]. These tasks have been tackled using various methods. For example, in [29], entity recognition was carried out using a Gated Recurrent Unit (GRU) network, and relation extraction was performed using probabilistic methods. In [30], the authors adopted a concept pruning strategy that is based on a measure called SemRefD to analyze references between concepts. This measure has been utilized in KGs to assess the likelihood of one concept being a prerequisite of another. In order to evaluate the relation extraction, the algorithm XGBoost was used for binary classification. In several other techniques, entity extraction was cast as a classification problem which was addressed using algorithms like Support Vector Machine (SVM) [31] and Conditional Random Field classifier (CRFClassifier) [32]. Other methodologies focused on entity extraction using semantic similarity [33] or Stack Long Short Term Memory (Stack-LSTM) with Spacy[1] [34]. In a similar way, some researchers used KGs for job market data representation. For example, the authors of [35] implemented entity linking using rule-based methods. Qin *et al.* [36] proposed a novel approach for entity recognition by feeding the word vectors into a bidirectional Long Short Term Memory and having the outputs fed into Conditional Random Fields layer (LSTM-CRF). They also improved the reliability of the extracted skill entities using click through data as a way for entity filtering. Additionally, the relation extraction process was done through discovering the hypernym-hyponym relations between skill entities.

B. REPRESENTATION LEARNING

Knowledge representation learning aims to map entities and relations of a KG onto a dense low-dimension feature space that captures the information and the structure of the graph. This was mainly carried out for job market-related graph applications. Studies such as [37] proposed representation learning algorithms such as Joint Bayesian Personalized Ranking (Joint-BPR) and Joint-Margin In [38], the authors proposed new algorithms for job representation learning using a directed graph as input. Other studies, such as [6], proposed a collective multi-view representation learning method by jointly learning topology view, semantic view, job transition balance view, and job transition duration view. In addition, they presented an encode-decode based fusion method to obtain a unified representation from the multi-view representations. JobBERT is the latest effort in this trajectory, where the authors of [39] presented a neural representation model for job titles, by augmenting a pre-trained language model with co-occurrence information from skill labels extracted from job vacancies.

Due to the literature gap in the area, using specific KG embedding techniques in educational KG construction was one of the direction [40] have undertaken. The authors presented a model for embedding learning of eduational KGs. Notably, the displayed method jointly learns embeddings from pre-trained structural (i.e. TransE) and literal embedding vectors (i.e. BERT). These are considered to be projections of KGs in the knowledge representation spaces e.g. Euclidean, non-Euclidean which is computationally efficient and representative but they suffer from lack of interpretability. This issue is remedied by rule based KG completion methods, which take advantage of the inherent structure of the KGs in order to find meaningful patterns.

What follows is a review of recent KG applications in education. Then, we present an analysis of KG based applications for employability, highlighting several factors that we think are critically important for a unified KG-based framework for education and the labor market.

IV. KG APPLICATIONS IN EDUCATION

Several studies have proposed methodologies to use KGs as a tool to improve education and its outcomes. We classify the reviewed studies into three main categories as shown in Figure 1.

A. ASSISTED INSTRUCTION

Assisted instruction refers to approaches to support instructors and faculty members in different aspects of their tasks. Several studies have addressed this area using KGs. In the next subsections, we describe two subcategories in which assisted instruction is the main topic.

1) INSTRUCTIONAL CONCEPTION

One of the main areas of educational research is instructional conception and its role in students learning. Instructional conception is the creation of learning experiences and materials in a manner that results in the acquisition and application of knowledge and skills. The authors of [29] conceptualized a system called KnowEDU that constructs a KG for education automatically. The result was a KG with nodes that represent instructional concepts that learners should master. The main purpose of KnowEDU is to identify the prerequisite relations between concepts in one course or subject, so that inter-course and inter-subject relations can be further explored. The instructional concepts are identified from pedagogical data using recurrent neural networks (RNN). The links between the instructional concepts are established

[1] https://spacy.io/universe/project/video-spacys-ner-model

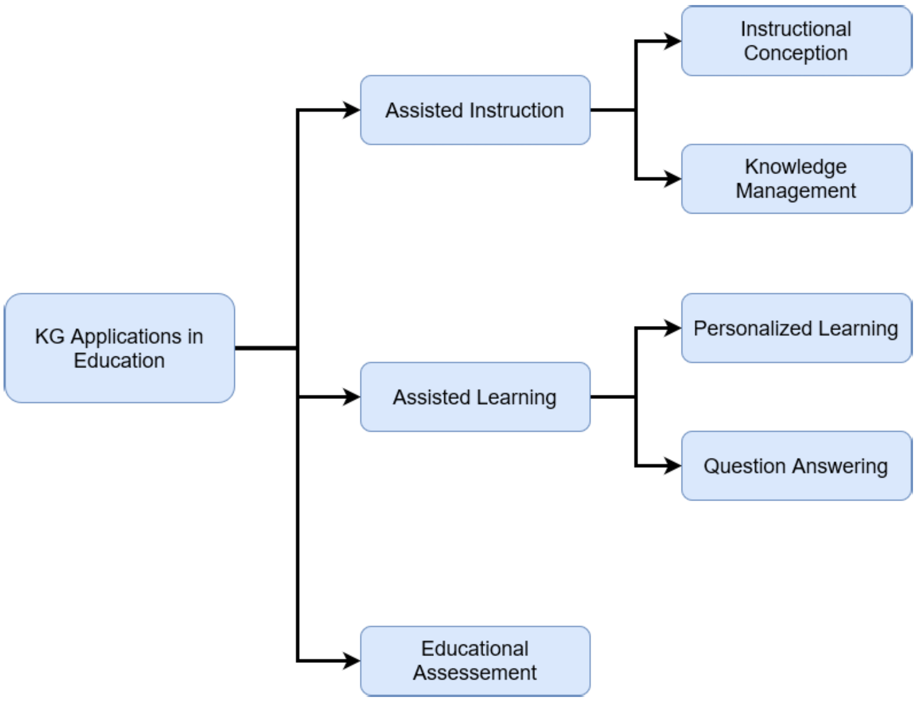

FIGURE 1. Taxonomy of KG applications in education.

through the probabilistic association rule mining algorithm by using students' performance data. The same methodology was used for the automatic construction of an educational KG from K-12 educational subjects [41]. In [42], the authors devised a procedure that enables the automatic construction of an educational KG using a model that combines Bidirectional Encoder Representations from Transformers (BERT), Bidirectional LSTM (BiLSTM) and the CRF model. Their KG was built using two large datasets, specifically subject teaching resources and an online encyclopedia.

[30] developed a strategy for identifying concepts' prerequisites based on the semantic relations found in a KG. The process consists of two phases: i) a candidate prerequisite searching module that generates an initial list of candidates, which are concepts that are members of the same category as the target concepts; the associations are established using properties like Common Memberships (CM), Linked Concepts (LC) in addition to novel pruning methods; and ii) a prerequisite evaluation module, that determines prerequisite relationships between concepts using a machine learning approach that models the problem as a binary classification. In [43], the authors presented a process to support domain experts in creating KGs by automating parts of their generation. They utilized a method that consisted of entity recognition, relation extraction and graph completion using the T-Mitocar tool. The study resulted in creating annotated educational KGs for personalized learning. In the same context, [33] adopted an acknowledged educational ontology to conceptualize and build a KG for Massive Open Online Courses. Entities represent 28591 instances with their corresponding relations, including 4 platforms, 604 universities, 18671 teachers and 9312 courses. [35] constructed a KG from video lectures by following a methodology that consists of transcript extraction from the video then applying named entity recognition, co-reference recognition and triple extraction. The proposed methodology however has many limitations such as language restrictions, time restrictions and accent restrictions.

2) KNOWLEDGE MANAGEMENT

Knowledge management using KGs explored tools such as SoLeMiO [31] to integrate learning material in popular authoring software The study focused on semantic representation by exploiting information from open KGs through expansion and filtering strategies in order to assist teachers in the authoring process of learning material through semi-automatic annotations. The authors used the identified concepts to extract and retrieve related learning resources from an open corpus. Furthermore, [44] presented a KG-based design of a scientific publication management model for scientific metadata using data analysis technologies. For evaluation, they implemented the use case of an entrepreneurship scientific publication management prototype system, in order to encourage sustainable learning. Similar work tackled courses allocation and management using KGs [45]. One of the latest studies in this direction of research was ELODIE [46], a mechanism for knowledge management. The study presented the task at hand as an information retrieval problem. The authors proposed a faceted search interface enhanced by a natural language query capabilities. ELODIE organises nodes and links of the KG by facets. Students can click on any provided option to formulate their queries. The user can then manipulate the dataset and visualizations according to their needs. The system presents a simple approach for engaging the user in information retrieval and knowledge management.

B. ASSISTED LEARNING

Researchers argued that using KGs in assessing students would improve educational outcomes. Based on our investigation, we group these applications into two subcategories, as described next.

- **Personalized Learning** Personalized learning in KGs was presented as techniques to help students engage and understand their study material. [47] constructed a personalized learning system based on knowledge structural graphs. They then defined and implemented an optimized learning path generator in order to provide learners with personalized learning content. Interactive applications using KGs is a relatively new direction to improve education. In [50], the authors addressed interactive visual navigation and inductive reasoning in KGs. Visualization was also the topic of the study carried out by [49]. The work evaluated the utility of the educational KG by a visualization analysis in order to provide a better comprehension of its structure.
- **Question Answering** [48] relied on crowdsourcing to enrich a math-related KG (MathGraph). The main purpose of the KG was to help students solve mathematical exercises with mathematical derivations and calculations. The authors also presented an algorithm to align a semantically parsed exercise to MathGraph in order to automatically resolve the mathematical problem.

C. EDUCATIONAL ASSESSMENT

Educational assessment focuses on assessment of a learner's knowledge and the level of skill they have acquired. Multiple studies tackle this aspect using KGs. [32] explored establishing teacher support methodologies where the authors adopted a KG-based approach to detect plagiarism. They focus on building a semantic KG from unstructured data. The built system implements entity recognition, followed by relationship extraction, then performs a co-reference resolution. It uses semantic similarity and sentence replacement for purpose of plagiarism detection.

KGs also contributed to advancing the question generation field especially in education. [51] implemented an end-to-end approach to generate natural language questions from knowledge graphs. An entity is first selected as an answer. A structured triple-pattern query having this answer as its sole result is then generated. If a multiple-choice question is desired, the approach selects alternative answer options. It also adopts a template based method to verbalize the structured query and yield a natural language question. The authors based their approach on an earlier methodology that follows the same pattern [52]. [53] presented a university faculty evaluation system based on a multi-view KG. The study started by constructing a KG that presents faculty information such as scientific research papers, patents, funds, monographs, awards, professional activities and teaching performance. Based on the KG embeddings of this data, an academic development factor (ADF) was proposed for making predictions about faculty academic development.

V. KG APPLICATIONS FOR EMPLOYABILITY

In previous studies, KGs are used to represent and visualize the entities and relationships related to social commerce, e.g. [54]. Inductive logic programming techniques were used to discover implicit information that can interpret the behaviors and intentions of the users. In this section, we explore more studies that focused on analyzing the job market using knowledge graphs.

A. JOB RECOMMENDATION

[38] tackled the job representation aspect in knowledge graphs. The authors developed a representation of job positions consisting of job title and company pairs which captures the similarity and the ordering relations among job positions. In the context of job recommendation, [37] created three types of information networks from the historical job data: job transition network, job-skill network, and skill co-occurrence network. The authors of [37] conceptualized a representation learning model that uses the information from all three networks. This study presented the pair wise ranking objective which learns job and skill vector representations into a shared latent space using three pre-processed graphs.

Recent studies started exploring the graph based recommendation systems. For example, [55] developed a recommender system by measuring the skill importance of each occupation for each country and feeding the best vector model to a graph database. [56] built a KG based recommender system for jobs called GUApp. The latter recommends relevant jobs to the users. The system also integrates a domain-specific KG that allows it to support natural dialogues with users about the discussed domain. This leads to an improvement in the quality of recommendations results. Last but not least, [57] proposed an explainable job recommendation system, which uses a KGs to model job-postings and user profiles into a unified framework. The system generates human-readable explanation for each recommendation based on the graph structure itself and a named entity classifier.

B. JOB-SKILL MATCHING AND TALENT INTELLIGENCE

The latest studies in this direction is the Skills & Occupation KG that was created in [58]. The KG was built by leveraging existing skills and occupation taxonomies enriched with external job posting. [35] built a KG representing the artificial intelligence job market and the corresponding skills required. The KG was used to explore the relationship of different key skills of artificial intelligence.

Talent intelligence exploration started with [6] presenting a data-driven approach for job title benchmarking. The authors

TABLE 1. KG applications in education.

Paper	Application Category	Approach	Methods	KG Resources
[19]	Instructional Conception	Courses to concepts mapping Concept-level dependencies learning Link prediction	CGL Class CGL Rank KNN	TEACHER dataset
[29]	Instructional Conception	Concept extraction Prerequisite relation identification	GRU Probabilistic association rule mining	The national curriculum of China
[30]	Instructional Conception	Candidate prerequisite searching Prerequisite evaluation	semRefD XGBoost	DBPedia, The CORE Corpus
[33]	Instructional Conception	Knowledge acquisition Knowledge fusion	Adhoc Semantic similarity	Coursera, EDX, XuetangX, Icourse websites
[35]	Instructional Conception	Named entity recognition Coreference recognition Relation extraction	Transition based Stack-LSTM Neural mention-ranking model Rule based approach	Video lectures
[42]	Instructional Conception	Knowledge points recognition Knowledge point association assessment	BERT-BiLSTM-CRF Semantic similarity and semantic relevance	Subject teaching resources Baidu Encyclopedia DBPedia
[31]	Knowledge Management	Data extraction Entity disambiguation and object alignment Onthology learning	SVM	DBpedia
[45]	Knowledge Management	Entity recognition Relation extraction and triple formation	Adhoc	Structural educational information system
[48]	Question Answering	Extracting operations and constraints Mapping entities	Crowdsourcing and domain experts	Complex, TriaNgle CoNiC Solid
[51]	Educational Assessment	Query generation Difficulty estimation	Logistic regression classifier	The Jeopardy! quiz-game show dataset
[32]	Educational Assessment	Entity recognition and relation extraction Sentence replacement	CRFClassifier Semantic similarity	ESCO Wikipedia

constructed a job title transition graph where nodes represent job titles and links denote the correlations between jobs. They purpose of the study was to provide an accurate guidance and considerable convenience for both talent recruitment and job seekers for position and salary calibration and prediction. Other like [36] constructed a KG of job skills through mining the abundant historical recruitment data. The authors utilized the constructed Skill-Graph for building a personalized question recommendation algorithm for improving job interview assessment.

A summary of the studies is provided in the reference tables 1 and 2.

VI. TOWARDS A UNIFIED FRAMEWORK FOR EDUCATION AND EMPLOYABILITY

Research has shown a lot of interest in diagnosing the issues that education and especially higher education suffers from, specifically when considering the job market. Previous work has analyzed the relationship between academia and the job market to understand how they influence each other [59]. [60] examined the collaborations between stakeholders and higher education institutions to identify the lack of communication between the job market and the youth. Other studies aimed to determine the relationship between skill mismatch and the regional unemployment. Researchers found that regions with higher skilled populations have lower unemployment rates [61]. [62] presented multiple data analyses and visualizations to understand discrepancies and temporal delays between evolving job market needs, course offerings, and science and technology developments. The findings of this work suggest that educational efforts play an important role in mediating between the needs of industry and research. [63] emphasized the importance of soft business-related skills and competencies and the need for prior work-experiences.

Investigating the dynamics between education and the job market is a very relevant issue which has been investigated in many ways before. Addressing this issue using KGs is an interesting research direction. We illustrate this using the following case study.

VII. CASE STUDY: DATA SCIENCE/AI

Here, we propose a framework to showcase the use of KG in the fusion of education and employability information. The framework is based on the idea that education provides students with certain skills and the job market needs certain skills to accomplish tasks in a certain job. This leads us to rely on skills to build links between the job market and education. Figure 3 illustrates a subgraph of the entire KG, built to describe the skills needed for a certain job and the courses that offer to teach those skills. We choose as a case study about data science jobs. We aim to give a full picture of what skills one would need in order to compete for a certain job. The schema should encompass all aspects of data science jobs, the skills needed for a given job and the courses that teach these skills. The KG may be used as recommender system for students to identify the skills needed for the job they want, for institutions to update their courses according to the job market evolution and eventually also to predict additional links.

We used a public dataset of data science/AI job offers collected from Indeed.[2] Each line in the dataset contains the job title, the company that posted the job offer and the

[2] www.indeed.com

TABLE 2. KG applications in employability.

Paper	Application Category	Approach	Methods	KG Resources
[37]	Job Recommendation	Construction of three graphs Graph embedding	Joint BPR Joint Margin	CareerBuilder plateform
[35]	Job-Skill Matching	Entity recognition Relationship extraction	LSTM	Online job offers
[38]	Job Recommendation	Job transition graph construction Node embeddings and Link prediction	Proposed algorithm for node embeddings Skipgram and APP	User resumes from the Indeed plateform
[55]	Job Recommendation	Word embeddings Graph database encoding	FastText Revealed comparative Advantage + Cosine similarity	Online Job vacancies ESCO dataset
[6]	Talent Intelligence	Job title aggregation Collective multi-representation Learning Link prediction	Fully connected layers	Online Professional Network
[2]	Talent Intelligence	Entity extraction Entity Linking Question recommendation	LSTM-CRF Gradient Boost Decision Tree Spearman's rank correlation coefficient	Job postings and candidates' resumes
[58]	Job-Skill Matching	Skill extraction Skill matching Link prediction	Textkernel Extract Jaccard similarity Node2Vec	ESCO and ISCO datasets Job postings from Jobdigger

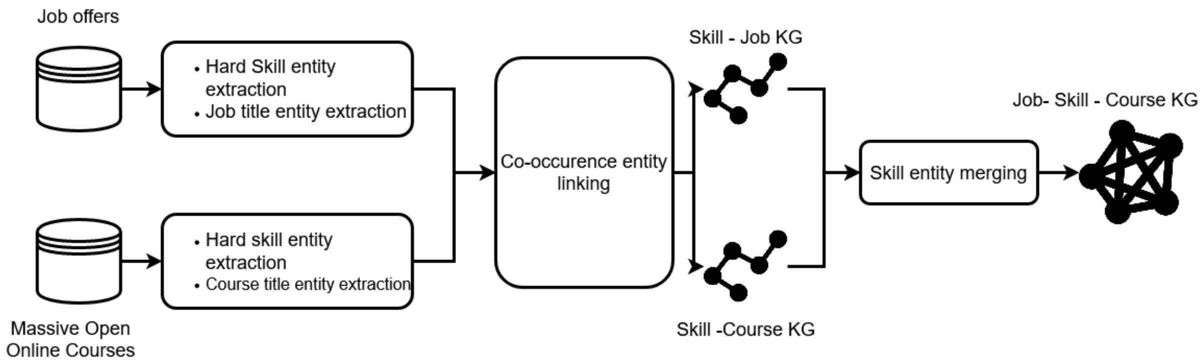

FIGURE 2. Methodology for job-skill-course KG construction.

hard skills needed for the job. We define hard skills as the technical skills needed to complete a certain task, whereas soft skills are abilities that are related to how an individual works and interacts with others. In our study, we focus solely on hard skills for simplicity reasons. We collected data about Massive Open Online Courses using a Web crawler where we gathered attributes such as the name of the course, the university that offers the course, and the language in which the course is taught. In addition, using the categories offered by the Website and entity recognition, we created a list of skills for each course, extracted from the course title, the description of the course and its category. A full view of the pipeline adopted in our framework is presented in Figure 2.

We created two weighted directed graphs: the first one has nodes as jobs and skill entities (Skill-Job KG) and the second one has nodes representing skills and courses (Skill-Course KG). In order to create the final KG, we rely on semantic matching for entity alignment between the Skill-Job KG and Skill-Course KG for the skill entity. The KG is heterogeneous as it contains three types of nodes: jobs, skills and courses. The job nodes are linked to certain skill nodes based on the skills needed for the job, and skill nodes are linked to certain course nodes based on the fact that these courses offer teaching those skills as part of their course goals. The weight of the relationship between nodes is the number of co-occurrences of these entities. We normalize the weights in order to have values between 0 and 1.

For visualization purposes, we chose the top ten skills required for a data scientist (the edges with the highest weights that are connected to the node 'data scientist'), as shown in Figure 3 and the top two courses for each skill (the edges with the highest weights that are connected to the chosen skills).

VIII. FUTURE RESEARCH DIRECTIONS

In this section, we discuss the limitations of the existing work and propose research directions for future work.

- **KG Construction and Data Quality** Most of the existing work paid very little attention to discussing the details of and the motivations behind the techniques used for entity recognition and relation extraction. It would be interesting to compare these techniques in order to identify those that perform the best in the different tasks. Furthermore, to the best of our knowledge, there are no methods to fully automate the process of KG construction. Developing such methods would be a promising research direction, particularly when dealing with large-scale unstructured content.

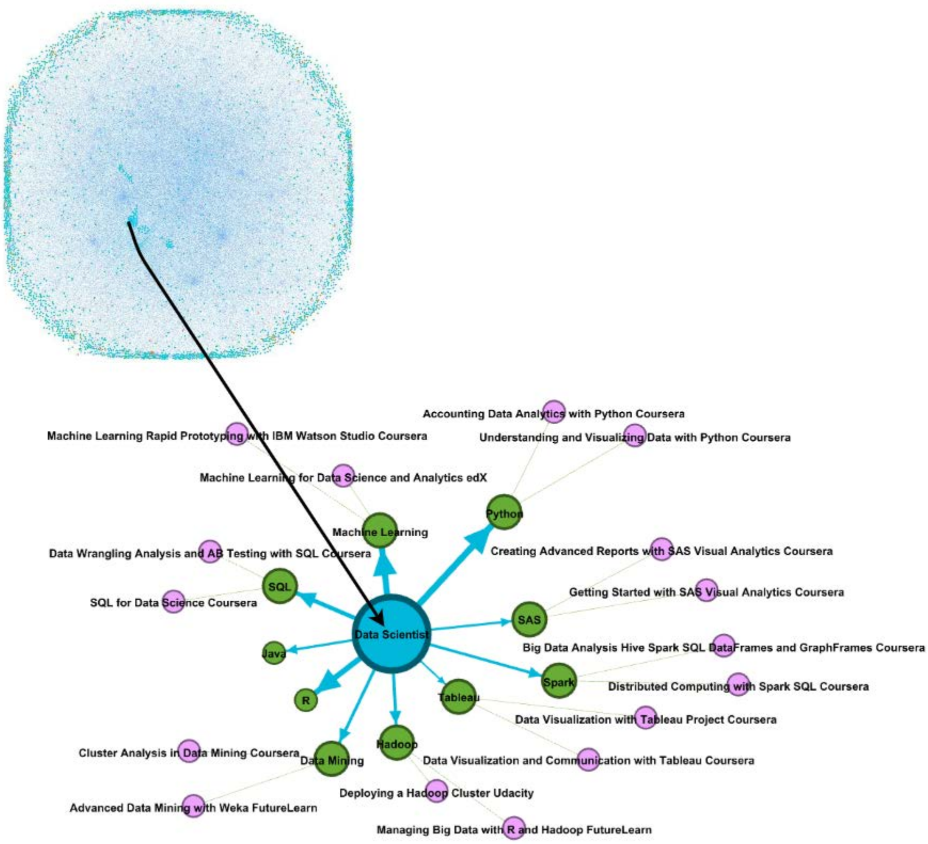

FIGURE 3. A snapshot of job-skill-course knowledge graph. The blue nodes represent the job titles, the pink nodes represent the skills, and the green nodes represent the courses.

The methodologies reviewed in this paper did not consider the use of standardized and appropriate data quality measures when constructing KGs. The data collection process gave little consideration to the credibility of the generated information. An interesting research direction would be to develop methods to evaluate the credibility of the information generated from KGs. This could be carried out by building probabilistic KGs where each relation/link is associated with a measure of its certainty.

- **Temporal KGs** Temporal KGs are one of the new directions that researchers have recently started to explore especially in education. This need to expansion comes from the fact that static KGs do not give a full representation about real world data and are only considered to be a static snapshot of multi-relational data. Studies such as [64] initiated research in this direction by constructing a novel deep evolutionary knowledge network that efficiently learns non-linearly evolving entity representations over time in a multi-relational setting.
  Work on incorporating temporal KGs in the job market is very limited. There has been some descriptive work done on using time-series on clusters of skillsets for job market forecasting [65], [66]. A dynamic KG, together with learning algorithms capturing dynamics, can address the limitation of traditional knowledge representation and reasoning by considering the temporal dimension [1].
- **KG-based recommendation** KG-based recommender systems attracted considerable interest in recent years. Unlike traditional recommender systems, KG-based recommender systems present a variety of entities and relations linking users and items, and illustrate better the reasoning process. Generating recommendations with the KG as side information for job recommendation and curriculum design would benefit institutions, job seekers and employers. Advancements in this field focus on multi-task learning where transferring knowledge from KG-related tasks, such as entity representation, classification and resolution would enhance recommendation performance [67].

IX. CONCLUSION

The use of knowledge graphs to improve education and employability has received much attention in recent literature. This study examined the different applications of such knowledge graphs as well as the approaches taken to construct them. We detailed these studies by categorizing them into different aspects of education and employability.

To contribute to reducing the mismatch between higher education and market need, we advocate the development and exploitation of knowledge graphs that jointly represent education and job market data. As a step towards this, we introduced a framework for building such knowledge graphs, and use the field of data science/AI as a case study. A number of research avenues remain to be explored, and this article discussed a few.

REFERENCES

[1] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A survey on knowledge graphs: Representation, acquisition and applications," 2020, *arXiv:2002.00388*.
[2] X. Zou, "A survey on application of knowledge graph," *J. Phys., Conf.*, vol. 1487, Mar. 2020, Art. no. 012016.
[3] M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, and S. Ali, "Using machine learning to predict Student difficulties from learning session data," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 381–407, Jun. 2019.
[4] F. Bousbahi and H. Chorfi, "MOOC-Rec: A case based recommender system for MOOCs," *Proc. Social Behav. Sci.*, vol. 195, pp. 1813–1822, Jul. 2015.
[5] A. Fedorets, F. Lottmann, and M. Stops, "Job matching in connected regional and occupational labour markets," *Regional Stud.*, vol. 53, no. 8, pp. 1085–1098, Aug. 2019.
[6] D. Zhang, J. Liu, H. Zhu, Y. Liu, L. Wang, P. Wang, and H. Xiong, "Job2Vec: Job title benchmarking with collective multi-view representation learning," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manage.*, 2019, pp. 2763–2771.
[7] I. Khaouja, G. Mezzour, K. M. Carley, and I. Kassou, "Building a soft skill taxonomy from job openings," *Social Netw. Anal. Mining*, vol. 9, no. 1, pp. 1–19, Dec. 2019.
[8] N. Pontika, "Roles and jobs in the open research scholarly communications environment: Analysing job descriptions to predict future trends," *LIBER Quarterly, J. Assoc. Eur. Res. Libraries*, vol. 29, no. 1, pp. 1–20, Jul. 2019.
[9] X. Dong, E. Gabrilovich, G. Heitz, W. Horn, N. Lao, K. Murphy, T. Strohmann, S. Sun, and W. Zhang, "Knowledge vault: A web-scale approach to probabilistic knowledge fusion," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2014, pp. 601–610.
[10] A. Hogan, E. Blomqvist, M. Cochez, C. d'Amato, G. de Melo, C. Gutierrez, J. E. L. Gayo, S. Kirrane, S. Neumaier, A. Polleres, R. Navigli, A.-C. N. Ngomo, S. M. Rashid, A. Rula, L. Schmelzeisen, J. Sequeda, S. Staab, and A. Zimmermann, "Knowledge graphs," 2020, *arXiv:2003.02320*.
[11] Y. Dai, S. Wang, N. N. Xiong, and W. Guo, "A survey on knowledge graph embedding: Approaches, applications and benchmarks," *Electronics*, vol. 9, no. 5, p. 750, May 2020.
[12] M. Papoutsoglou, A. Ampatzoglou, N. Mittas, and L. Angelis, "Extracting knowledge from on-line sources for software engineering labor market: A mapping study," *IEEE Access*, vol. 7, pp. 157595–157613, 2019.
[13] I. Khaouja, I. Kassou, and M. Ghogho, "A survey on skill identification from online job ads," *IEEE Access*, vol. 9, pp. 118134–118153, 2021.
[14] M. Rizun, "Knowledge graph application in education: A literature review," *Acta Universitatis Lodziensis. Folia Oeconomica*, vol. 3, no. 342, pp. 7–19, Aug. 2019.
[15] B. Abu-Salih, "Domain-specific knowledge graphs: A survey," *J. Netw. Comput. Appl.*, vol. 185, Jul. 2021, Art. no. 103076.
[16] C. Liang, J. Ye, Z. Wu, B. Pursel, and C. Giles, "Recovering concept prerequisite relations from university course dependencies," in *Proc. AAAI Conf. Artif. Intell.*, 2017, pp. 1–6.
[17] L. Pan, C. Li, J. Li, and J. Tang, "Prerequisite relation learning for concepts in moocs," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, vol. 1, 2017, pp. 1447–1456.
[18] J. Gordon, L. Zhu, A. Galstyan, P. Natarajan, and G. Burns, "Modeling concept dependencies in a scientific corpus," in *Proc. 54th Annu. Meeting Assoc. Comput. Linguistics*, 2016, pp. 866–875.
[19] Y. Yang, H. Liu, J. Carbonell, and W. Ma, "Concept graph learning from educational data," in *Proc. 8th ACM Int. Conf. Web Search Data Mining*, Feb. 2015, pp. 159–168.
[20] P. Talukdar and W. Cohen, "Crowdsourced comprehension: Predicting prerequisite structure in Wikipedia," in *Proc. 7th Workshop Building Educ. Appl. Using NLP*, pp. 307–315. 2012.
[21] S. Changuel, N. Labroche, and B. Bouchon-Meunier, "Resources sequencing using automatic prerequisite-outcome annotation," *ACM Trans. Intell. Syst. Technol.*, vol. 6, no. 1, pp. 1–30, Mar. 2015.
[22] F. Gasparetti, C. De Medio, C. Limongelli, F. Sciarrone, and M. Temperini, "Prerequisites between learning objects: Automatic extraction based on a machine learning approach," *Telematics Informat.*, vol. 35, no. 3, pp. 595–610, Jun. 2018.
[23] S. Wang, A. Ororbia, Z. Wu, K. Williams, C. Liang, B. Pursel, and C. L. Giles, "Using prerequisites to extract concept maps fromTextbooks," in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2016, pp. 317–326.
[24] C. Zhu, H. Zhu, H. Xiong, C. Ma, F. Xie, P. Ding, and P. Li, "Person-job fit: Adapting the right talent for the right job with joint representation learning," *ACM Trans. Manage. Inf. Syst.*, vol. 9, no. 3, pp. 1–17, Sep. 2018.
[25] C. Qin, H. Zhu, T. Xu, C. Zhu, C. Ma, E. Chen, and H. Xiong, "An enhanced neural network approach to person-job fit in talent recruitment," *ACM Trans. Inf. Syst.*, vol. 38, no. 2, pp. 1–33, Mar. 2020.
[26] W. Shalaby, B. Alaila, M. Korayem, L. Pournajaf, K. AlJadda, S. Quinn, and W. Zadrozny, "Help me find a job: A graph-based approach for job recommendation at scale," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2017, pp. 1544–1553.
[27] H. Xu, Z. Yu, J. Yang, H. Xiong, and H. Zhu, "Talent circle detection in job transition networks," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 655–664.
[28] Y. Mao, "Summary and evaluation of the application of knowledge graphs in education 2007–2020," *Discrete Dyn. Nature Soc.*, vol. 2021, pp. 1–10, Sep. 2021.
[29] P. Chen, Y. Lu, V. W. Zheng, X. Chen, and B. Yang, "KnowEdu: A system to construct knowledge graph for education," *IEEE Access*, vol. 6, pp. 31553–31563, 2018.
[30] R. Manrique, B. Pereira, and O. Mariño, "Exploring knowledge graphs for the identification of concept prerequisites," *Smart Learn. Environments*, vol. 6, no. 1, pp. 1–18, Dec. 2019.
[31] C. Grévisse, R. Manrique, O. Mariño, and S. Rothkugel, "Knowledge graph-based teacher support for learning material authoring," in *Proc. Colombian Conf. Comput.* Cham, Switzerland: Springer, 2018, pp. 177–191.
[32] K. Khadilkar, S. Kulkarni, and P. Bone, "Plagiarism detection using semantic knowledge graphs," in *Proc. 4th Int. Conf. Comput. Commun. Control Autom. (ICCUBEA)*, Aug. 2018, pp. 1–6.
[33] F. Dang, J. Tang, and S. Li, "MOOC-KG: A MOOC knowledge graph for cross-platform online learning resources," in *Proc. IEEE 9th Int. Conf. Electron. Inf. Emergency Commun. (ICEIEC)*, Jul. 2019, pp. 1–8.
[34] G. S. Shanmukhaa, S. K. Nandita, and M. V. K. Kiran, "Construction of knowledge graphs for video lectures," in *Proc. 6th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Mar. 2020, pp. 127–131.
[35] S. Jia, X. Liu, P. Zhao, C. Liu, L. Sun, and T. Peng, "Representation of job-skill in artificial intelligence with knowledge graph analysis," in *Proc. IEEE Symp. Product Compliance Eng. Asia (ISPCE-CN)*, Dec. 2018, pp. 1–6.
[36] C. Qin, H. Zhu, C. Zhu, T. Xu, F. Zhuang, C. Ma, J. Zhang, and H. Xiong, "DuerQuiz: A personalized question recommender system for intelligent job interview," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 2165–2173.
[37] V. S. Dave, B. Zhang, M. Al Hasan, K. AlJadda, and M. Korayem, "A combined representation learning approach for better job and skill recommendation," in *Proc. 27th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2018, pp. 1997–2005.
[38] H. Luo, S. Ma, A. J. B. Selvaraj, and Y. Sun, "Learning job representation using directed graph embedding," in *Proc. 1st Int. Workshop Deep Learn. Practice High-Dimensional Sparse Data*, 2019, pp. 1–5.
[39] J.-J. Decorte, J. Van Hautte, T. Demeester, and C. Develder, "JobBERT: Understanding job titles through skills," 2021, *arXiv:2109.09605*.
[40] S. Yao, R. Wang, S. Sun, D. Bu, and J. Liu, "Joint embedding learning of educational knowledge graphs," in *Artificial Intelligence Supported Educational Technologies*. Cham, Switzerland: Springer, 2020, pp. 209–224.
[41] P. Chen, Y. Lu, V. W. Zheng, X. Chen, and X. Li, "An automatic knowledge graph construction system for K-12 education," in *Proc. 5th Annu. ACM Conf. Learn. at Scale*, Jun. 2018, pp. 1–4.
[42] Y. Su and Y. Zhang, "Automatic construction of subject knowledge graph based on educational big data," in *Proc. 3rd Int. Conf. Big Data Educ.*, Apr. 2020, pp. 30–36.
[43] R. Manrique, B. Pereira, and O. Marino, "Exploring knowledge graphs for the identification of concept prerequisites," *Smart Learn. Environ.*, vol. 6, pp. 1–18, Dec. 2019.

[44] Y. Chi, Y. Qin, R. Song, and H. Xu, "Knowledge graph in smart education: A case study of entrepreneurship scientific publication management," *Sustainability*, vol. 10, no. 4, p. 995, Mar. 2018.

[45] I. Aliyu, A. F. D. Kana, and S. Aliyu, "Development of knowledge graph for university courses management," *Int. J. Educ. Manage. Eng.*, vol. 10, no. 2, p. 1, 2020.

[46] D. R. De, M. Garofalo, D. Malandrino, M. A. Pellegrino, and A. Petta, "Education meets knowledge graphs for the knowledge management," in *Proc. Int. Conf. Methodologies Intell. Syst. Technol. Enhanced Learn.* Cham, Switzerland: Springer, 2020, pp. 272–280.

[47] X. Huang, "Study of personalized E-learning system based on knowledge structural graph," *Proc. Eng.*, vol. 15, pp. 3366–3370, Jan. 2011.

[48] T. Zhao, C. Chai, Y. Luo, J. Feng, Y. Huang, S. Yang, H. Yuan, H. Li, K. Li, F. Zhu, and K. Pan, "Towards automatic mathematical exercise solving," *Data Sci. Eng.*, vol. 4, no. 3, pp. 179–192, Sep. 2019.

[49] K. Sun, Y. Liu, Z. Guo, and C. Wang, "EduVis: Visualization for education knowledge graph based on web data," in *Proc. 9th Int. Symp. Vis. Inf. Commun. Interact.*, Sep. 2016, pp. 138–139.

[50] V. Telnov and Y. Korovin, "Semantic web and interactive knowledge graphs as an educational technology," in *Cloud Computing Security-Concepts and Practice*. London, U.K.: IntechOpen, 2020.

[51] D. Seyler, M. Yahya, and K. Berberich, "Knowledge questions from knowledge graphs," in *Proc. ACM SIGIR Int. Conf. Theory Inf. Retr.*, Oct. 2017, pp. 11–18.

[52] D. Seyler, M. Yahya, and K. Berberich, "Generating quiz questions from knowledge graphs," in *Proc. 24th Int. Conf. World Wide Web*, May 2015, pp. 113–114.

[53] Q. Lin, Y. Zhu, H. Lu, K. Shi, and Z. Niu, "Improving university faculty evaluations via multi-view knowledge graph," *Future Gener. Comput. Syst.*, vol. 117, pp. 181–192, Apr. 2021.

[54] L. Yuan, Z. Huang, W. Zhao, and P. Stakhiyevich, "Interpreting and predicting social commerce intention based on knowledge graph analysis," *Electron. Commerce Res.*, vol. 20, no. 1, pp. 197–222, Mar. 2020.

[55] A. Giabelli, L. Malandri, F. Mercorio, M. Mezzanzanica, and A. Seveso, "Skills2Job: A recommender system that encodes job offer embeddings on graph databases," *Appl. Soft Comput.*, vol. 101, Mar. 2021, Art. no. 107049.

[56] G. M. Biancofiore, T. Di Noia, E. Di Sciascio, F. Narducci, and P. Pastore, "Enhancing job recommendations with knowledge graphs," MDPI, Basel, Switzerland, Tech. Rep. 116-138, 2021.

[57] C. Upadhyay, H. Abu-Rasheed, C. Weber, and M. Fathi, "Explainable job-posting recommendations using knowledge graphs and named entity recognition," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2021, pp. 3291–3296.

[58] M. de Groot, J. Schutte, and D. Graus, "Job posting-enriched knowledge graph for skills-based matching," 2021, *arXiv:2109.02554*.

[59] S. Angioni, A. A. Salatino, F. Osborne, D. R. Recupero, and E. Motta, "Integrating knowledge graphs for analysing academia and industry dynamics," in *Proc. ADBIS, TPDL EDA Common Workshops Doctoral Consortium*. Cham, Switzerland: Springer, 2020, pp. 219–225.

[60] I. Khaouja, I. Makdoun, G. Mezzour, I. Rahhal, H. Benchekroun, Y. El Hatib, and I. Kassou, "Social network analysis of job market stakeholders in Morocco," in *Proc. Int. Conf. Social Comput., Behav.-Cultural Modeling, Predict. Behav. Represent. Modeling Simulation*, Washington, DC, USA, 2017.

[61] A. Newell, "Skill mismatch and regional unemployment in Poland," in *The European Labour Market*. Cham, Switzerland: Springer, 2006, pp. 187–202.

[62] K. Börner, O. Scrivner, M. Gallant, S. Ma, X. Liu, K. Chewning, L. Wu, and J. A. Evans, "Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 50, pp. 12630–12637, Dec. 2018.

[63] J. Andrews and H. Higson, "Graduate employability, 'Soft skills' versus 'hard' business knowledge: A European study," *Higher Educ. Eur.*, vol. 33, no. 4, pp. 411–422, Dec. 2008.

[64] R. Trivedi, H. Dai, Y. Wang, and L. Song, "Know-evolve: Deep temporal reasoning for dynamic knowledge graphs," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 3462–3471.

[65] E. M. Sibarani and S. Scerri, "Generating an evolving skills network from job adverts for high-demand skillset discovery," in *Proc. Int. Conf. Web Inf. Syst. Eng.* Cham, Switzerland: Springer, 2020, pp. 441–457.

[66] E. M. Sibarani and S. Scerri, "SCoDis: Job advert-derived time series for high-demand skillset discovery and prediction," in *Proc. Int. Conf. Database Expert Syst. Appl.* Springer, 2020, pp. 366–381.

[67] 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. 2020.

YOUSRA FETTACH received the M.Sc. degree from ENSIAS, Rabat, Morocco, in 2019, with a major in data science and big data. She is currently pursuing the Ph.D. degree with the International University of Rabat. Her research interests include knowledge graphs and NLP. She was a recipient of the UIR research scholarship, in 2020.

MOUNIR GHOGHO received the M.Sc. and Ph.D. degrees from the National Polytechnic Institute of Toulouse, France, in 1993 and 1997, respectively. He was an EPSRC Research Fellow with the University of Strathclyde (Scotland), from September 1997 to November 2001. In December 2001, he joined the School of Electronic and Electrical Engineering, University of Leeds, U.K., where he was promoted to a Full Professor, in 2008. While still affiliated with the University of Leeds, in 2010, he joined the International University of Rabat (Morocco), where he is currently the Dean of the College of Doctoral Studies and the Director of the ICT Research Laboratory (TICLab). He was a recipient of the 2013 IBM Faculty Award, and a recipient of the 2000 U.K. Royal Academy of Engineering Research Fellowship. He is the Co-Founder and the Co-Director of the CNRS-Associated International Research Laboratory DataNet, in the field of big data and artificial intelligence. His research interests include machine learning, signal processing, and wireless communication, on which he has published over 300 papers in journals and conferences. He has coordinated around 20 research projects and supervised over 30 Ph.D. students in the U.K. and Morocco. In the past, he served as an Associate Editor for many journals, including the *IEEE Signal Processing Magazine* and the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

BOUALEM BENATALLAH (Fellow, IEEE) has been a Full Professor in computing with Dublin City University (DCU), Ireland, since January 2022. He has had over 21 years as a Senior Lecturer, an Associate Professor, a Professor, then a Scientia Professor with UNSW Sydney, Australia. His research interests include developing fundamental concepts and techniques in web service composition, services engineering, crowd sourcing services, data curation, cognitive services, and business processes management.