Journal of Innovation & Knowledge 7 (2022) 100190

Journal of Innovation & Knowledge

https://www.journals.elsevier.com/journal-of-innovation-and-knowledge

# Bridging the skill gap between the acquired university curriculum and the requirements of the job market: A data-driven analysis of scientific literature

Naif Radi Aljohani[a,*], Ahtisham Aslam[a], Alaa O. Khadidos[a], Saeed-Ul Hassan[b]

[a] Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
[b] Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, United Kingdom

A R T I C L E I N F O

*Article History:*
Received 27 December 2021
Accepted 6 April 2022
Available online 27 April 2022

Edited by Dr. J.R. PIÑEIRO

*Keywords:*
Skill gap
Acquired skill
Required skill
University curriculum
Job market
Big data analytics
Higher education
Upskill

A B S T R A C T

This research provides a comprehensive, first-of-its-kind, in-depth, data-driven analysis of the discussions on "curriculum alignment" in the light of "learned skills" and "acquired skills", as illustrated by cross-disciplinary records in Scopus. It was undertaken from 2010 to 2021 on 10,214 data points obtained to fully grasp the issues, names and themes that have contributed to the field over the past decade, and it presents the case for the increased value and new application of bibliometric analyses. When faced with scholarly research not included in Scopus, on the one hand, and links between previously divided research groups, on the other, ensuring the compatibility of the various scientific information archives is essential. Only in this manner can the research artefacts be made evident, the concerns and problems that have been either overlooked or under-researched be identified, and practical debate between academia and policymakers be facilitated.


## Introduction

Innovation and knowledge are the two main stumbling blocks to developing new technologies for society (Tiberius, Schwarzer & Roig-Dobón, 2021). Only by mastering the skills to mitigate the upcoming challenges and technological demands of Industry 5.0 (Ballestar de las Heras et al., 2020) can it be sure to develop innovation in the actual sprint (Hilmersson & Hilmersson, 2021). Thus, individuals must acquire the necessary information, skills, values and attitudes to live productive lives, make informed decisions and, through education, contribute both locally and internationally when facing and addressing global issues (Haddad, Haddad & Nagpal, 2021).

The rapid adoption of the Internet has undoubtedly revolutionized many aspects of our lives, from our patterns of interaction to our employment opportunities and how we search them out. This shift has resulted in a surge of online information, and it has become tricky to find the right facts (Jin, Wah, Cheng & Wang, 2015; Singh & Miah, 2020). The evolution of technology has also impacted on pedagogical and academic criteria, to the extent that curriculum designers need to be vigilant to provide students with abilities that meet the market's expectations (Kureková, Beblavý & Thum-Thysen, 2015; Vitale, Bowyer & Bayerlein, 2020).

The study of the job market is a growing field of interest, using new data sources and analytical tools, and is particularly important in practical fields such as computing (Woolridge, 2016). Keeping up with the employment market involves collecting, filtering and evaluating data from Internet job postings (Khaouja, Kassou & Ghogho, 2021; Smith & Ali, 2014). Generally, these 'postings' are seen as a valuable resource in analysing the skill sets required, and they have much potential in job market research. Posting jobs on the Internet started to gain momentum in the 1990s because it was faster and less expensive than putting out traditional job advertisements (Müller et al., 2016). This presented the prospect of accessing and analysing job postings to truly comprehend the industry's development in order to bridge the gap between the demands of academia and industry (Kim & Angnakoon, 2016), and it is seen as a tool to align the education curriculum to the market's requirements, gearing students up for employment.

The dynamic and agile nature of trends in both industry and the job market means that undertaking content analysis of information in online job listings manually is a chore that takes much time and

* Corresponding author.
*E-mail addresses:* nraljohani@kau.edu.sa (N.R. Aljohani), maaslam@kau.edu.sa (A. Aslam), aokhadidos@kau.edu.sa (A.O. Khadidos), s.ul-hassan@mmu.ac.uk (S.-U. Hassan).

https://doi.org/10.1016/j.jik.2022.100190

effort. Academics and curriculum developers must continually examine the employment market to keep their curricula (learned skills) parallel to these requirements (required skills). As the source of the skilled labour necessary to guarantee future prosperity (Aken et al., 2009), higher education plays a significant role in today's modern, post-industrial economy, yet students' successful transition into higher education and the job market is a long and complex process. Universities need to put greater emphasis on meeting students' expectations, assessing the needs of the target job market and modifying their offer to meet them. Semantic-based educational and analytical applications play a crucial role in the knowledge representation of academic courses, in modelling learning outcomes and in classifying jobs and their required skills, all of which can ultimately be used for the purposes of analysis and prediction (Wowczko, 2015).

Recently, many published scientific studies have highlighted this important and timely challenge of curriculum alignment to bridge the gap between the skills attained in the university and corresponding skills that are mandatory in the job market (Manaf, 2021; Zimmer & Keiper, 2021). To view the holistic research landscape, a comprehensive, data-driven perspective is desirable. To achieve this, our study used bibliometric analysis to investigate the scientific discussion and themes at the cross-section of "curriculum alignment", "acquired skills" and "required skills", arguing that bibliometric analysis has promise, value and application across multiple concerns and domains.

In view of the project's scope, its findings are inextricably linked to the significant issue of extensive data management and usage. In several academic fields, bibliometric indexes have been used statistically to assess the quality and impact of the literature in that discipline, making it more straightforward for researchers to review and analyse scientific papers (Guerrero-Bote, Moed & Moya-Anegón, 2021). Such investigations aim to uncover a field's academic structure by identifying the significant scholars and intellectual effects (Donthu et al., 2021). This research explores the potential of bibliometric analysis for studying the evolution of the issues in question. To achieve this goal, it uses state-of-the-art bibliometric indexes such as co-authorship networks, citation networks and keyword co-occurrence networks in a series of analyses to:

- compile a suitable set of scientific articles in the study space of "curriculum alignment" and associated research domains
- examine the development of the field of "curriculum alignment" statistically, through the lens of publication and citation counts
- determine the most powerful institutions and nations in the sector
- investigate the collaborative networks across institutions and authors
- determine the exchange of citations amongst source titles (conferences/journals)
- examine the field's chronological thematic progression
- query data-driven case studies in the realm of "curriculum alignment".

The remainder of the paper is structured to make a case for the field's evolution. Section 2′s discussion presents a problem-driven evaluation of literature relevant to the debate on the subject. The data method is discussed in Section 3, which details the bibliometric analysis and associated keywords. Section 4 presents the results of the bibliometric indexes. Section 5 summarizes and closes the case by listing the additional study opportunities made feasible by this work. The additional value and unique ability of bibliometric analysis to track the profession's growth are argued throughout and, as a result, new sectors of bibliometric analytic application are described.

## Literature review

There has been much research into the various strategies for analysing the employment market, from manual analysis to data mining (Eom & Lim, 2012; Todd, McKeen & Gallupe, 1995). A recent report discusses the problem of student retention and drop out, observing that one student in three either fails or drops out soon after enrolment (Müller et al., 2016). Non-completion not only influences individual students' personal and professional life but has financial implications for their family, the university and government. Moreover, it affects society and the economy through the loss of important skills and knowledge. There are financial and reputational implications for universities around the world, making student retention and success issues of concern. Researchers have proposed many theoretical models to explain the underlying concepts and reasons for student attrition, student satisfaction and retention in higher education. Watson et al., (Watson, Young, Miranda, Robichaux & Seerley, 1990) studied alumni behaviour and investigated their satisfaction level, proposing a satisfaction model to assist institutes to identify areas of improvement and formulate strategies for continuous progress and development. They found that multiple factors impacted on student satisfaction, course design being of the utmost importance, and inferred that courses should be designed to match real-world market trends and opportunities. Gorgone et al. (Gorgone et al., 2003) stress the importance of embedding both academic and employability skills into courses, and propose a skillset method incorporating these throughout the curriculum.

Concept mapping and the graphic representation of ideas and topics help students to understand course learning outcomes (CLO). Similarly, skills representation and mapping of course topics can help with student satisfaction and planning learning strategies. The concept mapping approach in education and semantic-based knowledge representation is presented by Gorgone et al. (Gorgone et al., 2005), who propose a theoretical model for semantic-based curriculum planning and assessment of students' understanding of concepts and topics. Its results can be used by students to plan their professional futures better, contributing to improved retention and employability.

In addition to rapidly revolutionizing learning trends and objectives in education, the field of information technology has made available vast amounts of online data on jobs. Finding and classifying the required facts in this huge volume of information is a complex task (Lee, 2005). It can be most helpful to use a supervised classification learning algorithm/method, in which a computer program trains and learns from an available dataset and uses this to classify new datasets. Applying ontological classification and reasoning to the domain data adds a further layer to supervised learning, yielding better results (Sodhi & Son, 2010). It is helpful in labelling job advertisements alongside the required and acquired skills parameters, and ultimately may identify any gap between a student's skills and the job market's requirements.

Due to the diversified trends in the job market and the rapid advances in the technology sector, it has become more difficult than ever to link university curricula to the job market. Indeed, graduates are typically sceptical at the start, leading them to waste time and resources in searching for jobs that are not right for them. Employers are looking for graduates with subject-specific experience and competencies (Zhang, Li & Zhang, 2012), an a disparity is apparent whereby recent start-ups and small companies are inclined to hire fresh graduates, while companies with a mature infrastructure prefer experienced employees. To maximize their available resources, mature companies are hesitant to choose to train new graduates to develop their employability skills (Gorgone et al., 2003), instead aiming to attract and hire skilled graduates. Therefore, there is a need for an employer-driven approach to minimize the gulf between the degree curriculum and industry's needs. From the very first year of a course, universities need to put special emphasis on career counselling and job paths so that students think clearly about moving ahead and focus on a career path that is aligned to their studies.

For some time, the research community has emphasized frameworks and multiple techniques to help students to acquire the appropriate computing skills. It started in the 1960s, with classical techniques such as surveys, questionnaires, one-to-one interviews and meetings with technology stakeholders to compile the technical skills to meet the market's demands (Aken et al., 2009; Albin & Otto, 1987; Todd et al., 1995; Watson et al., 1990). Later, in the mid-1990s, the research community designed tools to assess on a regular basis the skills needed, as well as to bring curricula into line with market expectations. 'Bridging the skills gap' is the term used for this procedure (Aken et al., 2009; Eom & Lim, 2012; Gorgone et al., 2003, 2005). There is a substantial literature on the strategies to do so, using fairly comparable methods that gather from multiple online sources the data relating to a profession (Aken et al., 2009; Debortoli, Müller & Vom Brocke, 2014; Kim & Lee, 2016; Lee, 2005; Smith & Ali, 2014; Sodhi & Son, 2010; Woolridge & Parks, 2016; Wowczko, 2015; Zhang et al., 2012), either manually, through web crawlers or from newspapers (Todd et al., 1995). Next, they categorize the occupational needs using analytical techniques (Lee, 2005; Müller et al., 2016; Todd et al., 1995), empirical methods (Aken et al., 2009; Sodhi & Son, 2010; Woolridge & Parks, 2016; Wowczko, 2015) or both (Kim & Lee, 2016). Furthermore, to perform manual content analysis (Lee, 2005; Todd et al., 1995; Woolridge & Parks, 2016) or automated analysis, data-mining techniques (clustering, classification, LSA) (Aken et al., 2009; Müller et al., 2016; Smith & Ali, 2014; Wowczko, 2015; Zhang et al., 2012) and other means (Kim & Lee, 2016; Sodhi & Son, 2010) are used on these categories in the datasets. A study by Woolridge and Parks (Woolridge & Parks, 2016) employed a hybrid approach, comparing online job vacancies and their postings to the curriculum in that area by a manual content process. As indicated by the literature, academics, educators and others are concerned to minimize any discrepancy between the skills demanded by industry and those possessed by graduates. A significant number of research studies have used manual analysis to extract the varied proportions of skills mentioned in job postings. This constraint has motivated researchers to bridge the gap between curricula and job skills by proposing new technological approaches and theoretical frameworks to assist decision-makers to formulate optimal policies for the identification of skilled labour (Westergaard, Stærfeldt, Tønsberg, Jensen & Brunak, 2018).

More recently, several studies have been published to stress the need to bridge the skill gap between university curricula and job market requirements and, more explicitly, to keep within sight of Industrial 4.0 and Industrial 5.0 needs in a diverse society. Oraison, Konjarski and Howe (Oraison et al., 2019) argue that, according to analyses of employability criteria, employers prioritize graduates with practical competences and 21st-century abilities such as problem-solving and communication. However, they barely makes mention of cultural understandings and attitudes toward inclusion and diversity, both of which are fundamental traits in graduates, likewise the need for professional accreditation. Therefore, the role of universities in preparing their students for an increasingly diverse modern society in contemporary times has become vital.

More recently, a plan known as the multiple-tier co-op model proposes that employers should select students as early as high school so that they can learn the skills relevant to their employer's needs at the same time as participating in a related educational programme (Hoanca & Craig, 2019). While the model has had a significant impact on the IT sector, the wide-scale implications of such a model for Industry 4.0 and Industry 5.0 are yet to be explored. Nevertheless, despite greater automation, the focus of Industry 4.0 is on humans, placing new demands on employees through new technologies, organisational structures and work processes. The digital transformation is a period of change that involves a high degree of adaptability and the fulfilment of new requirements (Brezeanu & Lazarou, 2020).

More recent literature in this field has highlighted the needs by proposing models such as product-based learning via the alignment of competencies with relevant industries (Yudiono et al., 2021), 'importance-performance analysis' of graduates' competencies gap relative to industry (Febriani et al., 2021) and vocational school alignment, with forthcoming industrial challenges (Mahmudah & Santosa, 2021). Finally, the application of artificial intelligence and machine learning to bridge the gap between university teaching and required skills is found in the more recent literature (Arora & Kohli, 2022; Bakkar & Axmann, 2022; Bigan, 2021).

## Data and methodology

Bibliographic analysis serves to determine the geographies, societies, institutions and countries contributing most proactively to a discipline (Hassan, Haddawy, Kuinkel, Degelsegger & Blasy, 2012), as well as to reveal the cutting-edge trends and collaboration patterns amongst research groups (Ferreira et al., 2016; Leydesdorff & Rafols, 2009). Various bibliographic indicators such as co-authorship maps, citation networks and word frequency maps are important in this type of analysis, and a detailed discussion is provided below. These metrics were applied to the data corpus to investigate and study the research landscape of curriculum alignment and related topics.

### *Bibliography dataset for the study*

In bibliography databases, scientific articles are organized into pre-set categories. For instance, Scopus uses the All-Science Journal Classification (ASJC) to organize sources (journals/conference proceedings) into a hierarchical structure of fields and subfields. Similarly, to organize sources systematically the Thomson Reuters ISI Web of Science uses subject categories (Bonaccorsi, Cicero, Haddawy & Hassan, 2017; Hassan, Haddawy & Zhu, 2014, 2016; Zhu, Yang, Xie, Wang & Hassan, 2014). In multidisciplinary domains, however, no single hierarchical sorting strategy can accommodate the entire publishing dataset (Hassan, Visvizi & Waheed, 2019; Safder & Hassan, 2019; Safder et al., 2020; Sarwar & Hassan, 2015; Waheed, Hassan, Aljohani & Wasif, 2018). As a result, in this study whole sets of keywords peculiar to curriculum alignment and skills gap were used to match the titles, abstracts and author keywords of publications indexed in the Scopus database. With the help of domain experts, the irrelevant publications were removed from the retrieved source titles. Using this method, we created a Scopus-compatible query for all terms linked to curriculum alignment, and then extracted the associated scientific publications. The term "publication" refers to any scientific writing, such as a review or scientific paper, that has been published in a recognized technical publication. The following is the Scopus-compatible query used in the study: title-abs-key ("curriculum" and "job*") or title-abs-key ("curriculum alignment") or title-abs-key ("education" and "students" and "unemployment") or title-abs-key ("university" and "unemployment") or title-abs-key ("skills gap") or title-abs-key ("job postings") or title-abs-key ("acquired skills") or title-abs-key ("required skills") and (limit-to (pubyear, 2021) or limit-to (pubyear, 2020) or limit-to (pubyear, 2019) or limit-to (pubyear, 2018) or limit-to (pubyear, 2017) or limit-to (pubyear, 2016) or limit-to (pubyear, 2015) or limit-to (pubyear, 2014) or limit-to (pubyear, 2013) or limit-to (pubyear, 2012) or limit-to (pubyear, 2011) or limit-to (pubyear, 2010)

Scopus provided the data in .csv (comma separated values) format. The files contain bibliographic data such as publication name and year, citation count, affiliation and source title, all of which assist bibliometric analysis. A total of 10,214 papers were obtained from 1 January 2010 to 3 November 2021. The extracted data were pre-

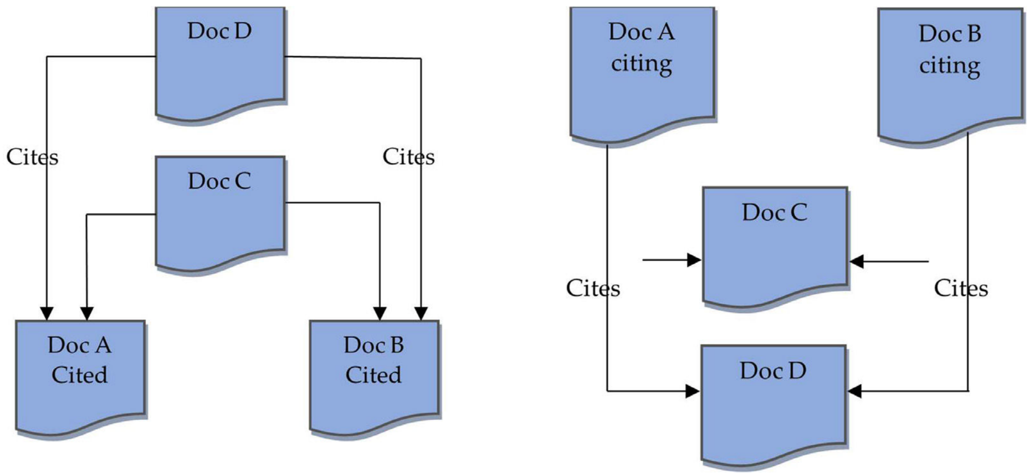


**Fig. 1.** Co-citation network and bibliography coupling.

processed by removing both non-essential details, such as the editor and sponsor, and entries with no author name. The extracted dataset contains articles, conference papers, book chapters and reviews in these numbers: articles (6582), conference papers (2305), book chapters (541), reviews (382), notes (121), books (92), editorials (72), short surveys (65) and letters (Aljohani, Fayoumi & Hassan, 2021).

The following attributes were included in our bibliometric analysis of the field of curriculum alignment: author, publication title, year of publication, source title, cited by, affiliation, author with affiliation, author keyword, index keyword, language, references and electronic identification of document (EID). To identify the regions of the world active in curriculum alignment research, the country names were derived from the affiliation.

*Bibliometric indicators*

The bibliometric measures applied to the procured dataset are described in this section, which presents the details of each bibliometric indicator.

*Co-authorship network*

The degree of collaboration amongst authors is measured by their co-authorship network (Aljohani et al., 2021; Grodzinski, Grodzinski & Davies, 2021). Scientific collaborations facilitate the construction of social networks, the improvement and expansion of knowledge horizons, and fostering of cognitive development (Huang, Gui, Qin & Du, 2021). Collaboration is most common in multidisciplinary domains, in which authors from multiple fields strive to create an intellectual research result that defines the field's knowledge structure (Hassan & Haddawy, 2013). These networks facilitate the interpretation of scientists' behavioural and communicative features in numerous disciplines, emphasizing the dynamics of knowledge flows. Scientific articles published by a collaboration of multiple researchers identify the interaction between institutes and give an overview of the flow of information (Hassan & Haddawy, 2015). In this study, we compiled the association networks amongst authors and their affiliated countries to depict the associations graphically. The details of various types of networks are provided below.

*Citation network*

In bibliometrics, citation networks are constructed via scientific publications through co-citation techniques. They assist in measuring the dependencies between various entities in the dataset, such as between scientists, journals, institutions and countries. Co-citation reflects the relativity of two articles when they independently cite one or two similar articles (Hota, Subramanian & Narayanamurthy, 2020). A scenario of co-citation is depicted in Fig. 1, where a research article (Document C) cites another study (Document A). It reflects their association, referred to as co-citation. A similar phenomenon is bibliographic coupling, where two articles (Document A, B) cite a common reference/study (Document D) (Cai, Chen, Huang & Ritter, 2021; Ferreira, 2018). This implies that they work in relatively similar research areas. This study presents co-citation maps and bibliographic associations with respect to source titles, and the details are provided in subsequent sections.

*Term co-occurrence network*

To study in depth the structure of curriculum alignment and skills, term maps were constructed using VOSviewer version 1.6.16 to explore the progressive development of this field and identify recent research themes and trends in its underlying field. This enabled us to explore techniques for data-driven decision-making in this discipline.

To create networks from the VOSviewer utility, terms from each article's title, abstract and author-defined keywords were used as the input, henceforth referred to as candidate terms. From these designated candidate terms VOSviewer constructed conceptual networks on the recommended normalization method 1, with 0 convergence, 1 resolution and 100 random starts. Each candidate term extracted from the title/abstract/author-defined keywords is represented as a node in the conceptual network, and the size of the node denotes the frequency of that term in the article. Nodes are positioned in the network on the basis of their associated strength: if two terms are strongly linked and similar, their nodes are close to each other (Van Eck & Waltman, 2007, 2010). This association and similarity of the terms can be represented by Equation 1.

$$s_{ij} = \frac{c_{ij}}{w_i w_j} \tag{1}$$

In the above equation, $s_{ij}$ refers to the connected similarity, $c_{ij}$ denotes the amount of co-occurrence between two terms *i* and *j*, $w_i$ is the individual number of occurrences of the term *i* and $w_j$ the individual number of occurrences of the term *j*. These conceptual networks in VOSviewer assist in visualizing the research streams of the terms used extensively in this discipline's research landscape.

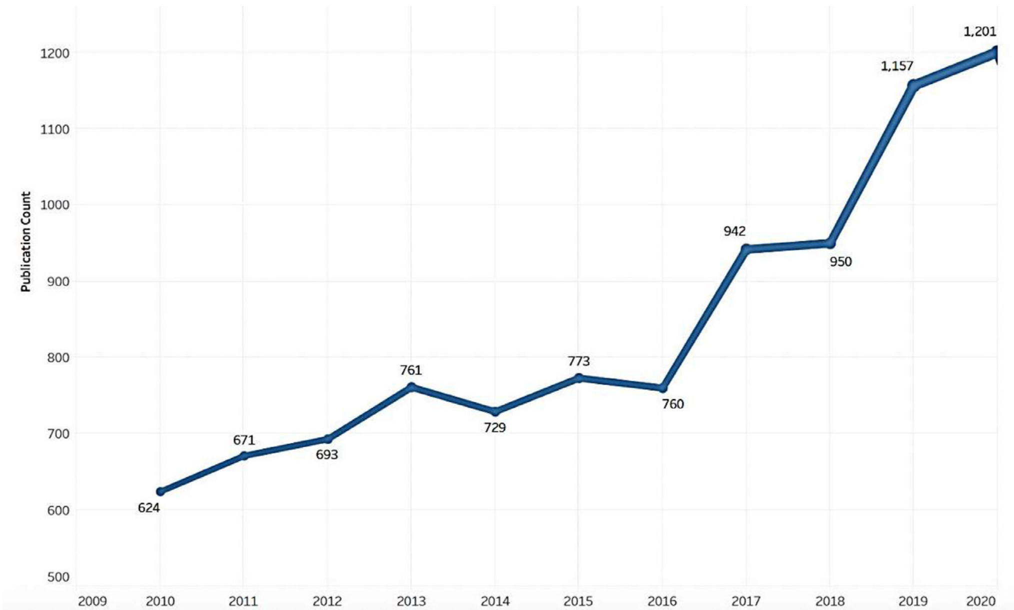


**Fig. 2.** Publication output in the field, 2010−2021.

## Results and discussion

In this section an in-depth analysis of the bibliometric analysis conducted on the curriculum alignment discipline is first presented, then discussed, in terms of the number of publications, research active countries and institutes, their citations and collaborations from 2010 to 2021. This study makes a holistic analysis of this discipline, analysing the research activities of prevalent countries, their number of publications and the citations of each publication per country. The country data were extracted from the affiliations section, collected from the Scopus corpus. These country-wise and citations-per-publication analyses were presented visually using Tableau Desktop[1] (free student version). For a more thorough and qualitative analysis of the corpus of this field, term co-occurrence maps were constructed using VOSviewer. A detailed explanation of these analyses is presented in the following subsections.

### *Discussion on publication output of countries and institutions*

This section presents the research landscape of this discipline in terms of the number of publications in each year by country. As depicted in Fig. 2, the number of publications has increased each year, reflecting the activity and the community's growing research interest in this field. To perform a more in-depth exploration of this discipline, keyword analysis was undertaken in Tableau software to reveal the evolution of the keywords prevalent in the research community. As seen in Fig. 3, prominent keywords are represented on the basis of their occurrence each year. Also, the total occurrence is provided, showing that 'COVID-19′ mentions only emerged in 2020. Around 55 occurrences of this keyword were seen then, and it appeared in greater numbers in 2021 (in bold). Similarly, the keyword 'Industry 4.0′ first emerged in 2017, and by 2020 it had gained rapid recognition in terms of its occurrence (in bold).

Table 1 shows the top 10 countries active in this research community with respect to publication counts. It can be seen that the United States is ahead of any other country by a large margin, making it the most active worldwide. Similarly, countries' output is depicted in Fig. 4(a) by a map showing publication count. Fig. 4(b) depicts citations per publication (CPP) for each country by shading, where lower values are light and higher values darker. This compares a country's citation count to its publication count. It can be observed that although some countries have a low publication count, their CPP ratio may be higher than one with a high publication count. For instance, the United States has the highest publication count, yet its CPP is lower than that of Canada, Australia, the United Kingdom or Sweden. Similarly, although Sweden has a low publication count it has the highest CPP, depicted in Fig. 4(b) by the darker shade.

Table 2 lists the top 10 institutions active in the discipline of curriculum alignment. The University of Toronto has the highest publication count, with 71 publications, followed by the University of Oxford, with a count of 49. Though the United States is the top country by publication count, by institution it is third place after Canada and the United Kingdom. This highlights that although Canada and the United Kingdom have a lower publication count than the United States, their institutions' publication counts surpass those of the United States. This also reveals that only a few institutions in these countries are active in this research area.

### *Discussion on collaboration network amongst countries and authors*

Collaboration networks depict the association of multiple entities to produce a research output. In this study, collaboration networks of authors and countries were constructed to determine the associations amongst them. Furthermore, the publication and citation counts of each first author (in the Scopus corpus) are depicted in Fig. 5 (left). Interestingly, although Stones and Zhang are the top two authors, producing the most publications, they have attracted little or no citation. On the other hand, although authors Stefanidis and Lee have few publications, just six and four respectively, they are cited the most: 371 and 400 citations, respectively. This demonstrates that both authors not only produce quality work but have their own research communities, as shown in the collaboration network in Fig. 4 (right). Also, from the collaboration network it can be observed that the group associated with Stones is very small, implying only limited participation in terms of numbers of citations.

Fig. 6 presents the co-authorship network amongst countries from 2010 to 2021. It can be observed in Fig. 6 (left) that clusters of countries are formed, showing countries' collaboration with each other. Each is represented in a different colour and reveals strong connections. Scandinavian countries form a cluster with Austria, France and Ukraine, similarly a prominent cluster is formed of the United States

[1] https://www.tableau.com/

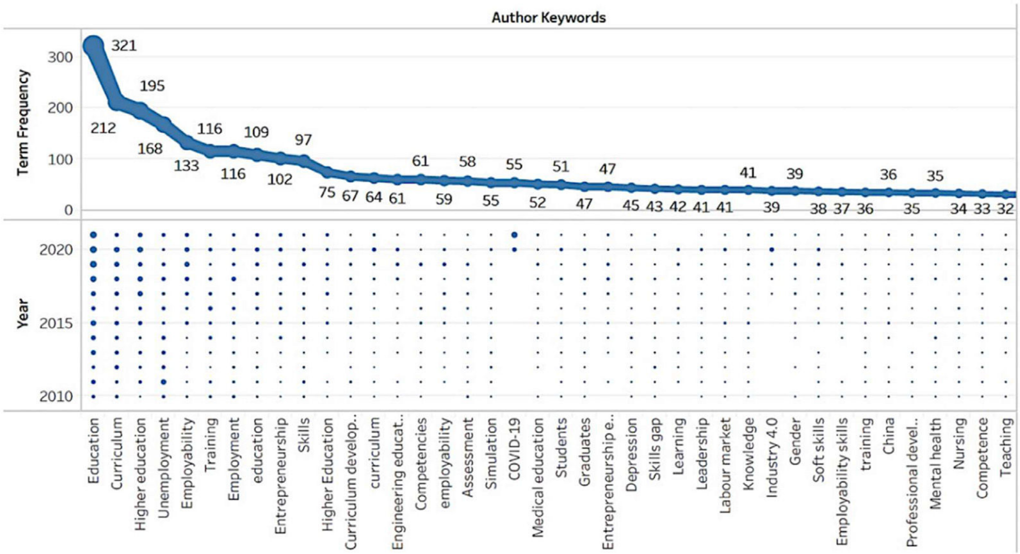


**Fig. 3.** Trends in author keywords in the field, by number of records and year, 2010–2021.

**Table 1**
Top 10 countries in the dataset by publication count, 2010–2021.

| Country | Publication count |
|---|---|
| United States | 3217 |
| United Kingdom | 942 |
| Australia | 551 |
| Germany | 501 |
| Canada | 415 |
| India | 348 |
| Spain | 331 |
| China | 328 |
| South Africa | 297 |
| Italy | 243 |

**Table 2**
Top 10 institutions in the dataset by publication count, 2010–2021.

| Institution | Publication count |
|---|---|
| University of Toronto | 71 |
| University of Oxford | 49 |
| University of Michigan, Ann arbour | 49 |
| Purdue University | 47 |
| University of Melbourne | 47 |
| Harvard Medical School | 45 |
| Monash University | 45 |
| Universiteit Maastricht | 43 |
| University of North Carolina at Chapel Hill | 43 |
| University of Minnesota Twin Cities | 43 |

with India, Taiwan, Nigeria, the United Kingdom, and so on. Saudi Arabia has a strong collaboration with Pakistan, Jordan, Egypt, and so on. Fig. 6 (left) shows these clusters by year. Recent collaborations are depicted in neon yellow, while older ones are darker. A recent collaboration pattern can be seen amongst Saudi Arabia, Morocco, Ukraine, Chile, Vietnam, and so on, countries that have started to network in this field.

*Discussion on citation network amongst sources*

For a more in-depth analysis of this research landscape, a network of sources that cite one another was constructed, using VOSviewer to identify influential sources in this field, as depicted in Fig. 7. The network shows the top venues, taking a minimum of five documents per source as the threshold by which to identify key source titles in the field. The size of the clusters represents the number of citations received by each source. It can be seen that a few major clusters are formed of sources that cite one another, such as medical education, including journals on surgery and nurse education. Similarly, there is a cluster for health, disorders, autism, and so on, and another is formed around economics. It can be seen that there is no prominent source with a network to every source: the sources form multiple clusters. Also, the count of journals is more numerous than that of conference papers. More on these journals and their studies is discussed in the next section.

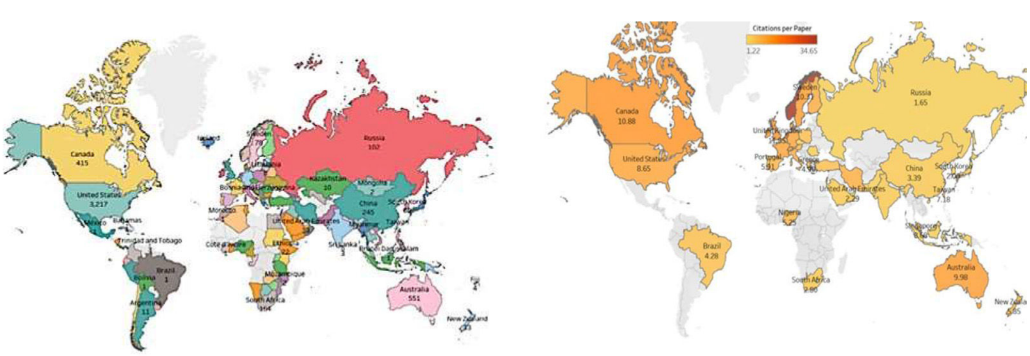


**Fig. 4.** Publication and citation counts of countries: (a) Publication count; (b) Citations per publication.

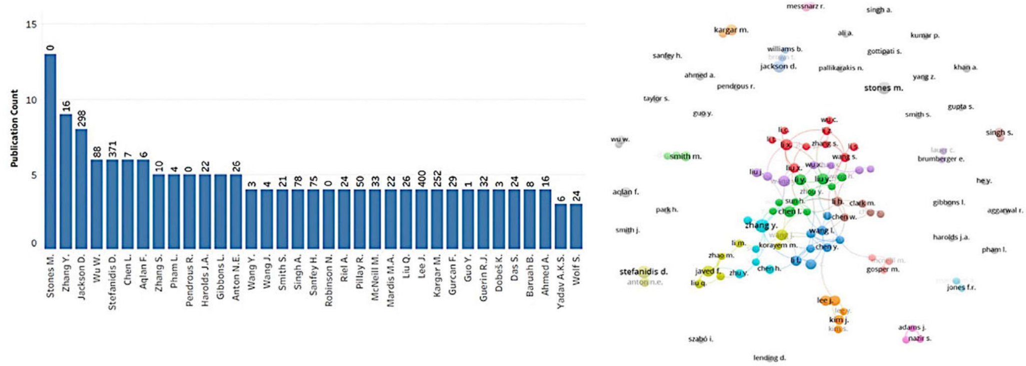


**Fig. 5.** Publication/citation counts of authors and co-authorship networks, 2010–2021.

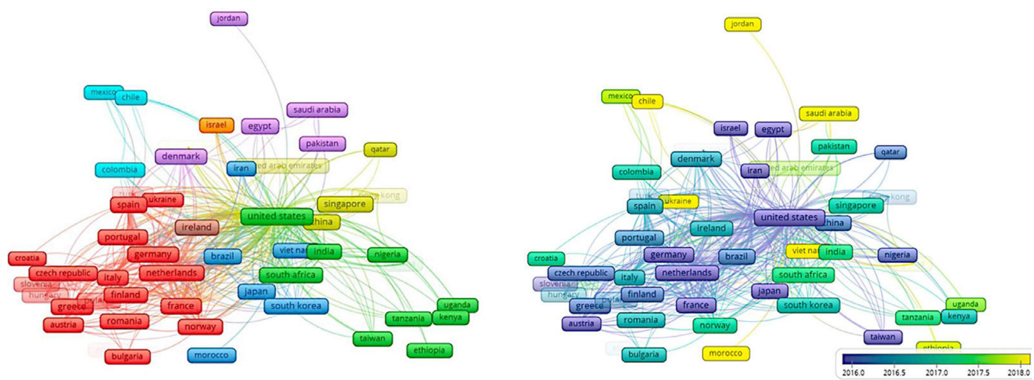

**Fig. 6.** Co-authorship network amongst nations, 2010–2021.

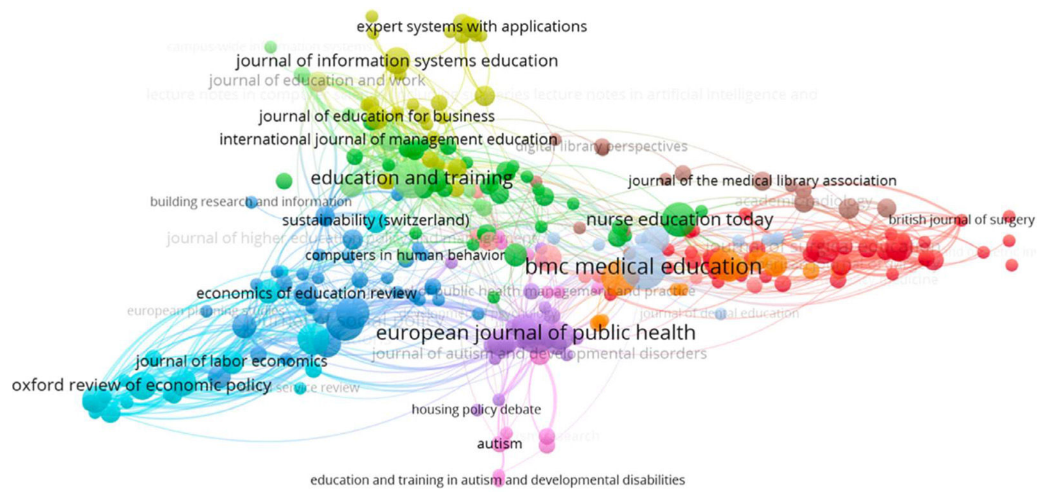


**Fig. 7.** Bibliographic coupling amongst sources, 2010–2021.

*Discussion on co-occurrence author-keyword network*

To examine the predominant keywords in this field, a text-based map was constructed in VOSviewer from the author-defined keywords from the Scopus corpus, as depicted in Fig. 8. The map shows 955 terms out of 19,433, using a minimum five occurrences of a keyword as the threshold by which to identify prominent research themes in a field. Related research themes cluster together in a single colour, revealing several research streams. The larger the cluster, the greater the number of the relevant research themes that are grouped in it. Some prominent research themes identified by the clusters are education, curriculum, higher education, employability, unemployment, training, university, gender and skills.

Drilling down to individual clusters, it is apparent that employment and unemployment are associated with mental health, depression, suicide and similar psychological tendencies. A further cluster consists of education, training, medical education, surgery, nursing, residency and patient safety, all deeply rooted in an individual's skills. The cluster of curriculum is formed with machine learning, blended learning, big data, collaborative learning, team formation, and so on.

Fig. 9 presents the most recent themes in this discipline, depicted in neon yellow. It can be observed that in terms of curriculum alignment and skills the most recent keywords to emerge are cybersecurity, cybersecurity education, machine learning, big data, collaborative learning, well-being, digitalization, Saudi Arabia, Mexico and

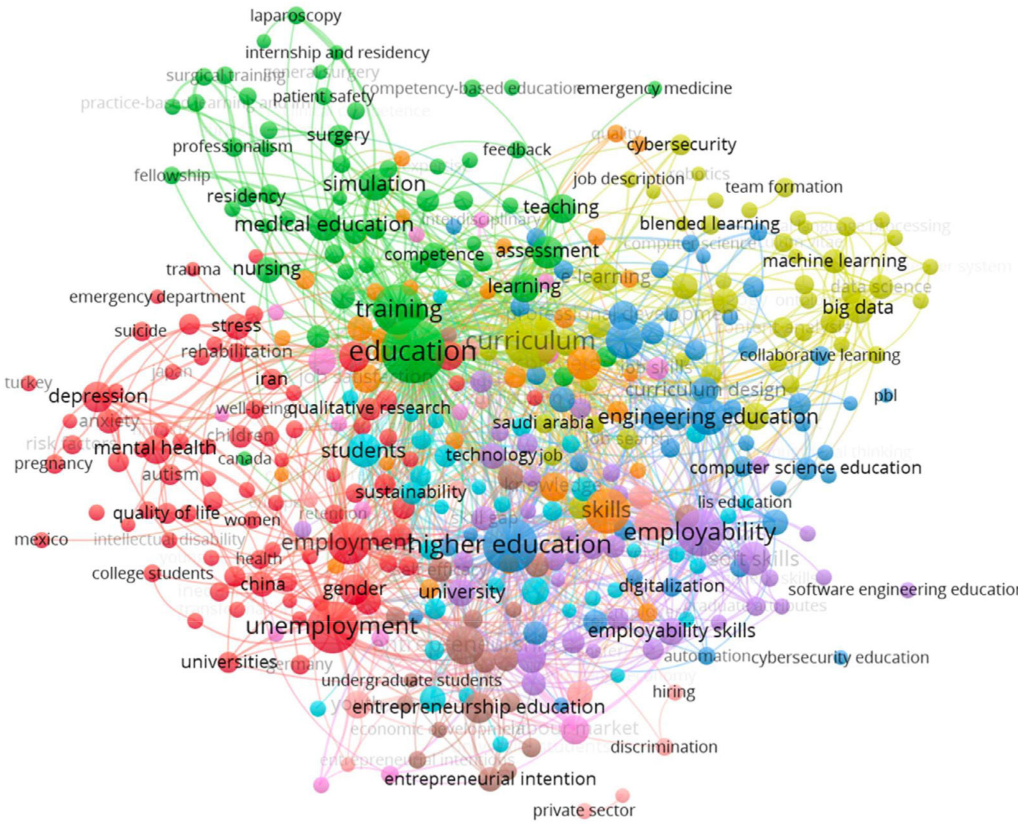


**Fig. 8.** Co-occurrence network of authored-defined keywords, 2010–2021.

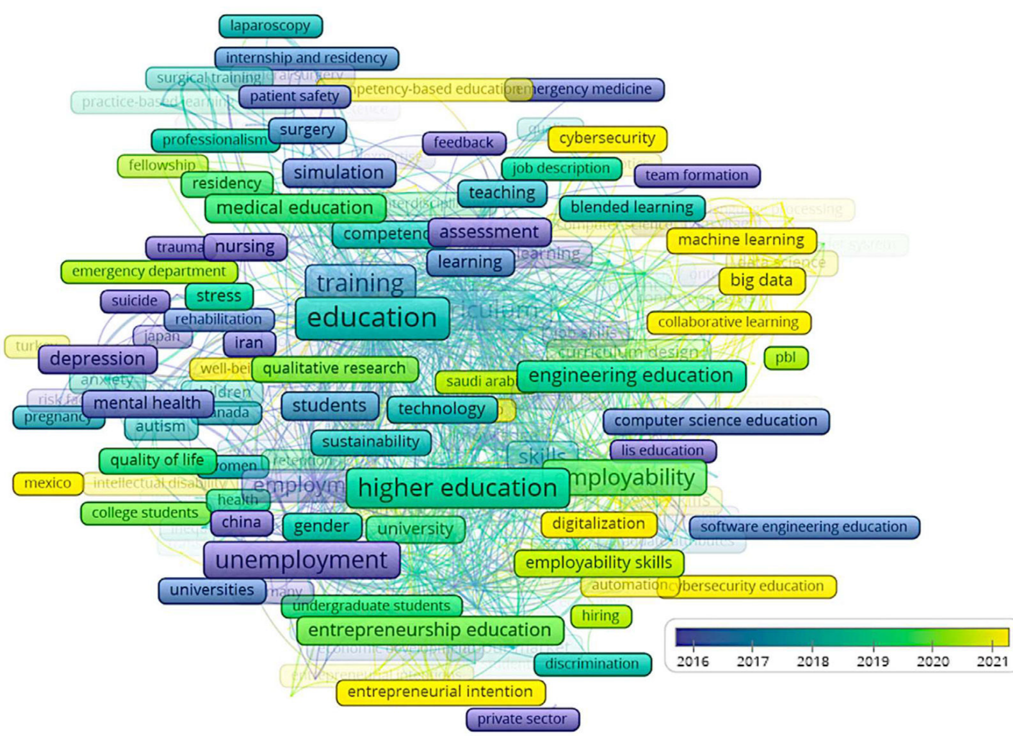


**Fig. 9.** Analysing the thematic evolution of the literature over time.

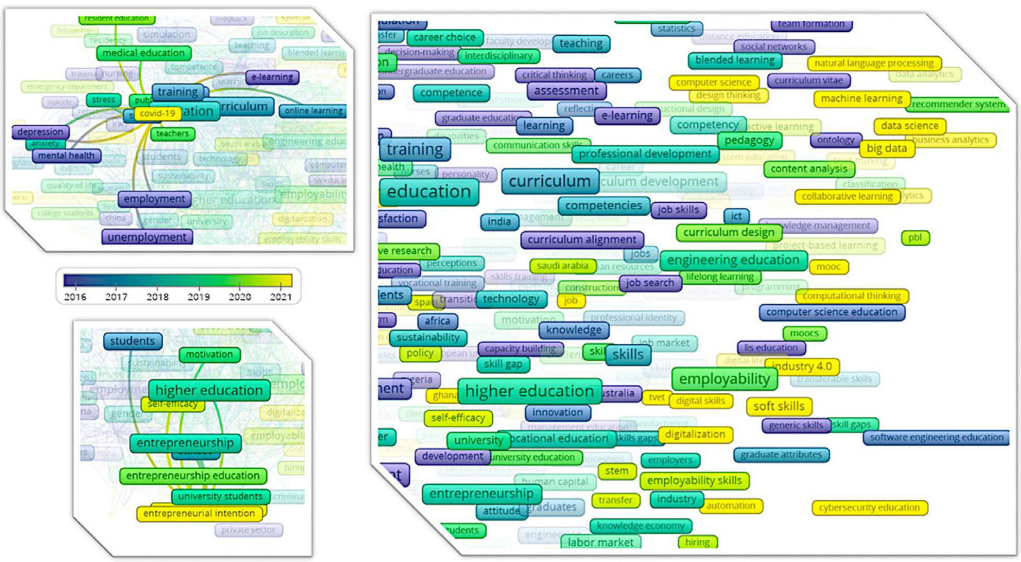


**Fig. 10.** Themes emerging from the literature in recent years.

'entrepreneurial intention'. These represent the recent themes to have garnered most attention from the research community.

For a more thorough analysis, from this keyword network we drilled down from selected keywords. In Fig. 10, the recent term 'covid-19′ is displayed in a large size, portraying its associated terms linked to unemployment, medical education, mental health, anxiety, depression, stress, curriculum, online learning, e-learning and teachers. These keywords cohesively reflect Covid's effects and how it impacts on society. Drilling down from 'higher education' reveals that this is linked to entrepreneurship, entrepreneurship education and intention, motivation, students and university. This cluster shows the significance of a university education in building the skills of entrepreneurship and students' ambition to start up an enterprise.

*Recommendations for policymakers*

Since the nature of professions is constantly changing due to technological breakthroughs and innovation, businesses need to make substantial efforts to train their employees. We show that artificial intelligence, blockchain, machine learning, cyber security, the internet of things and data visualisation are just a few new technologies that are transforming businesses. Thus, tuning into these technologies to stay relevant is now vital, not just for businesses but for students. To mitigate recent industrial challenges, these emerging technical skills must become part of an updated curriculum.

Technical skills demonstrate mastery and knowledge of specific, measurable talent, whereas soft skills frequently indicate working with people and growing within an organisation. Nevertheless, soft skills aid in building relationships and the resolution of problems, allowing the technical ability of individuals to be fully utilized. Also, soft skills demonstrate that an individual can comprehend the many characteristics that would allow them to thrive at an organisation in a given role. Our study shows that universities exhibit a clear gap in training students to acquire such skills. Employers recognize the importance of these skills in the workplace, accordingly universities must prioritize soft skills in the curriculum.

## Concluding remarks and limitations

This study presents a bibliometric analysis of the discipline of curriculum alignment to explore its research landscape and significance in the light of 'acquired' and 'learned' skills. The discussion in this study references the article corpus extracted from Scopus databases from 2010 to 2021, revealing research productivity and active research themes, countries and authors. The study is the first of its kind, presenting a detailed historic analysis of this discipline and its research.

Unsurprisingly, the United States emerged as the top country in terms of numbers of publications in this discipline, highlighting its research productivity. It surpasses other countries by a wide margin. In terms of citations per paper, Sweden is the country at the top, reflecting the quality of its research in this area. Although the number of its publications is far fewer, Sweden's citations per publication are high. The same contradictory pattern is seen for authors, as those with a high number of publications have few or no citations and those with a low publication count have a high citation count. Several clusters are observed in the author network, reflecting authors' community formation, which also impacts on their citation. No discrete pattern can be traced in countries' collaboration. Several cluster with others, demonstrating that they form research communities and work together in groups.

Keyword-based thematic analysis reveals the influential terms to have emerged in recent times, including but not limited to Covid-19 and the psychological issues pertaining to it, such as depression, anxiety, suicide, employability issues and concerns. These show the impact of Covid on education and associated mental states. While entrepreneurship education and entrepreneurial intention appear as prominent themes amongst university students, digital skills and soft skills are coupled with the Industry 4.0 revolution. Overall, in recent times the skills in high demand relate to data science, cybersecurity, natural language processing and machine learning.

## Declaration of Competing Intrest

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

## Funding

This research work was funded by Institutional Fund Projects under grant no. (IFPRC-104-611-2020). Therefore, authors gratefully acknowledge technical and financial support from the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Kingdom of Saudi Arabia.

## References


Aken, A., et al. (2009). Mining for computing jobs. *IEEE Software*, *27*(1), 78–85.

Albin, M., & Otto, R. W. (1987). The CIS curriculum: What employers want from CIS and general business majors. *Journal of Computing and Information System*, *27*, 15–19.

Aljohani, N. R., Fayoumi, A., & Hassan, S.-U. (2021). A novel deep neural network-based approach to measure scholarly research dissemination using citation networks. *Applied Science*, *11*(22), 10970.

Arora, Y., & Kohli, R. (2022). A study for making universities intelligent using data analytics and artificial intelligence. *Intelligent system algorithms and applications in science and technology* (pp. 147−155). Abingdon: CRC Press.

Bakkar, M. N., & Axmann, M. (2022). Industry 4.0: Learning analytics using artificial intelligence and advanced industry applications. In E. McKay (Ed.), *Manage your own learning analytics* (pp. 193−204). Springer Ed..

Ballestar de las Heras, M. T., et al. (2020). Productivity and employment effects of digital complementarities. *Journal of Innovation & Knowledge* preprint.

Bigan, C. (2021). Trends in teaching artificial intelligence for Industry 5.0. In A. Draghici, & L. Ivascu (Eds.), *Sustainability and innovation in manufacturing enterprises* (pp. 257−274). Springer Eds.

Bonaccorsi, A., Cicero, T., Haddawy, P., & Hassan, S.-U. (2017). Explaining the transatlantic gap in research excellence. *Scientometrics*, *110*(1), 217–241.

Brezeanu, T.-M., & Lazarou, E. (2020). Alignment Between Engineering Curriculum and skills Development for Industry 4.0. In Proceedings of International Scientific Conference eLearning and Software for Education (pp. 328−334) April 30 − May 1.

Cai, T., Chen, C., Huang, T. H., & Ritter, F. E. (2021). What Makes a Good Reference Manager? A quantitative analysis of bibliography management applications. In Proceedings of the Asian CHI Symposium (pp. 64−69) May 8−13.

Debortoli, S., Müller, O., & Vom Brocke, J. (2014). Comparing business intelligence and big data skills: A text mining study using job advertisements. *Business & Information Systems Engineering*, *6*, 289–300.

Donthu, N., et al. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, *133*, 285–296.

Eom, M., & Lim, C. (2012). Critical skills to be competent and relevant IT personnel: Do today's IT personnel possess requisite skills. *Journal of Information amd Technology Management*, *23*(4), 33–49.

Febriani, R. A., et al. (2021). Gap analysis of graduates' competencies in manufacturing engineering department with the industry. *IJMESH*, *4*(2), 75–84.

Ferreira, F. A. (2018). Mapping the field of arts-based management: Bibliographic coupling and co-citation analyses. *Journal of Business Research*, *85*, 348–357.

Ferreira, J. M., Ferreira, F. A. F., Fernandes, C. M., Jalali, M. S., Raposo, M. L., & Marques, C. S. (2016). What do we [not] know about technology entrepreneurship research? *International Entrepreneurship Management Journal*, *12*(3), 713–733.

Gorgone, J. T., Davis, G. B., Valacich, J. S., Topi, H., Feinstein, D. L., & Longenecker, H. E., Jr. (2003). IS 2002 Model curriculum and guidelines for undergraduate degree programs in information systems. *Communications of the Association for Information Systems*, *11*, 63.

Gorgone, J. T., Gray, P., Stohr, E. A., Valacich, J. S., & Wigand, R. T. (2005). MSIS 2006 curriculum preview. *Communications of the Association for Information Systems*, *15*, 544–554.

Grodzinski, N., Grodzinski, B., & Davies, B. M. (2021). Can co-authorship networks be used to predict author research impact? A machine-learning based analysis within the field of degenerative cervical myelopathy research. *Plos One*, *16*,(9) e0256997.

Guerrero-Bote, V. P., Moed, H. F., & Moya-Anegón, F. (2021). New indicators of the technological impact of scientific production. *Journal of Data Information Science.*, *6*(4), 36–61.

Haddad, G., Haddad, G., & Nagpal, G. (2021). Can students' perception of the diverse learning environment affect their intentions toward entrepreneurship? *Journal of Innovation Knowledge*, *6*(3), 167–176.

Hassan, S.-U., & Haddawy, P. (2013). Measuring international knowledge flows and scholarly impact of scientific research. *Scientometrics*, *94*(1), 163–179.

Hassan, S.-U., & Haddawy, P. (2015). Analyzing knowledge flows of scientific literature through semantic links: A case study in the field of energy. *Scientometrics*, *103*(1), 33–46.

Hassan, S.-U., Haddawy, P., Kuinkel, P., Degelsegger, A., & Blasy, C. (2012). A bibliometric study of research activity in ASEAN related to the EU in FP7 priority areas. *Scientometrics*, *91*(3), 1035–1051.

Hassan, S.-U., Haddawy, P., & Zhu, J. (2014). A bibliometric study of the world's research activity in sustainable development and its sub-areas using scientific literature. *Scientometrics*, *99*(2), 549–579.

Hassan, S. U., Sarwar, R., & Muazzam, A. (2016). Tapping into intra-and international collaborations of the Organization of Islamic Cooperation states across science and technology disciplines. *Science & Public Policy*, *43*(5), 690–701.

Hassan, S.-U., Visvizi, A., & Waheed, H. (2019). The 'who' and the 'what 'in international migration research: Data-driven analysis of Scopus-indexed scientific literature. *Behaviour and Information Technology*, *38*(9), 924–939.

Hilmersson, F. P., & Hilmersson, M. (2021). Networking to accelerate the pace of SME innovations. *Journal of Innovation & Knowledge*, *6*(1), 43–49.

Hoanca, B., & Craig, B. (2019). Building a K-16-industry partnership to train IT professionals. *Journal of Information System Education*, *30*(4), 232.

Hota, P. K., Subramanian, B., & Narayanamurthy, G. (2020). Mapping the intellectual structure of social entrepreneurship research: A citation/co-citation analysis. *Journal of Business Ethics*, *166*(1), 89–114.

Huang, X., Gui, Q., Qin, X., & Du, D. (2021). Spatial evolution of the global scientific collaborative network. *Regional Studies*, *8*(1), 127–129.

Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*, *2*, 59–64.

Khaouja, I., Kassou, I., & Ghogho, M. (2021). A survey on skill identification from online job ads. *IEEE access : practical innovations, open solutions*, *9*, 118134–118153.

Kim, J., & Angnakoon, P. (2016). Research using job advertisements: A methodological assessment. *Library & Information Science Research.*, *38*, 327–335.

Kim, J. Y., & Lee, C. K. (2016). An empirical analysis of requirements for data scientists using online job postings. *International Journal of Software Engineering and its Application*, *10*, 161–172.

Kureková, L. M., Beblavý, M., & Thum-Thysen, A. (2015). Using online vacancies and web surveys to analyse the labour market: A methodological inquiry. *IZA Journal Labor Economics*, *4*, 2.

Lee, C. K. (2005). Analysis of skill requirements for systems analysts in Fortune 500 organizations. *Journal of Computer Information System*, *45*, 84–92.

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of Association Information*, *60*(2), 348–362.

Mahmudah, F. N., & Santosa, B. (2021). Vocational school alignment based on industry needs. *JOVES*, *4*(1), 36–45.

Manaf, P. A. (2021). Inter-organizational alliance in improving the quality of learning through academic and industry alignment. *Cometra Education*, *1*(1).

Müller, O., Schmiedel, T., Gorbacheva, E., Brocke, J., Müller, O., Schmiedel, T., ... Brocke, J. (2016). Towards a typology of business process management professionals: Identifying patterns of competences through latent semantic analysis. *Enterprise Information Systems*, *10*, 50–80.

Oraison, H., Konjarski, L., & Howe, S. (2019). Does university prepare students for employment?: Alignment between graduate attributes, accreditation requirements and industry employability criteria. *JTLGE* preprint.

Safder, I., & Hassan, S.-U. (2019). Bibliometric-enhanced information retrieval: A novel deep feature engineering approach for algorithm searching from full-text publications. *Scientometrics*, *119*(1), 257–277.

Safder, I., Hassan, S.-U., Visvizi, A., Noraset, T., Nawaz, R., & Tuarob, S. (2020). Deep learning-based extraction of algorithmic metadata in full-text scholarly documents. *Information processing & management*, *57*,(6) 102269. doi:10.1016/j.ipm.2020.102269.

Sarwar, R., & Hassan, S.-U. (2015). A bibliometric assessment of scientific productivity and international collaboration of the Islamic world in science and technology (S&T) areas. *Scientometrics*, *105*(2), 1059–1077.

Singh, H., & Miah, S. J. (2020). Smart education literature: A theoretical analysis. *Education and Information Technology*, *25*(4), 3299–3328.

Smith, D., & Ali, A. (2014). Analyzing computer programming job trend using web data mining. Literature review − Web data mining. *Issues Information Science and Information Technology*, *11*, 203–214.

Sodhi, M. S., & Son, B.-G. (2010). Content analysis of OR job advertisements to infer required skills. *Journal of the Operational Research Society*, *61*, 1315–1327.

Tiberius, V., Schwarzer, H., & Roig-Dobón, S. (2021). Radical innovations: Between established knowledge and future research opportunities. *Journal of Innovation Knowledge*, *6*(3), 145–153.

Todd, P. A., McKeen, J. D., & Gallupe, R. B. (1995). The evolution of IS job skills: A content analysis of IS job advertisements from 1970 to 1990. *Mis Q.*, *19*, 1–27.

Van Eck, N. J., & Waltman, L. (2007). VOS: A New Method for Visualizing Similarities Between Objects. Eds, In H.-J. Lenz, & R. Decker (Eds.), *Advances in Data Analysis: Proceedings of the 30th Annual Conference of the German Classification Society* Eds. (pp. 299−306). Berlin, Germany: Springer.

Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538 2010.

Vitale, C., Bowyer, D., & Bayerlein, L. (2020). Developing and presenting a framework for meeting industry, student and educator expectations in university degrees. *e-Journal of Business Education and Scholarship of Teaching*, *14*(1), 57–65.

Waheed, H., Hassan, S.-U., Aljohani, N. R., & Wasif, M. (2018). A bibliometric perspective of learning analytics research landscape. *Behaviour and Information Technology*, *37*(10−11), 941–957.

Watson, H. J., Young, D., Miranda, S., Robichaux, B., & Seerley, R. (1990). Requisite skills for new MIS hires. *ACM Sigmis Database*, *21*, 20–29.

Westergaard, D., Stærfeldt, H.-H., Tønsberg, C., Jensen, L. J., & Brunak, S. (2018). A comprehensive and quantitative comparison of text-mining in 15 million full-text articles versus their corresponding abstracts. *PLoS computational biology*, *14*, e1005962.

Woolridge, R. W. (2016). What's in and what's out: Defining an industry-aligned IS curriculum using job advertisements. Rachida Parks University of Arkansas at Little Rock. *Journal of Higher Education Theory Practice*, *16*, 105–120.

Woolridge, R. W., & Parks, R. (2016). What's in and what's out: Defining an industry-aligned IS curriculum using job advertisements. *JHET*, *16*(2), 105.

Wowczko, I. (2015). Skills and vacancy analysis with data mining techniques. *Informatics*, *2*, 31–49.

Yudiono, H., et al. (2021). Product-based learning model through the alignment of mechanical engineering competencies with industry. *Jurnal Pendidikan Teknologi dan Kejuruan*, *27*(1).

Zhang, S., Li, H., & Zhang, S. (2012). Job opportunity finding by text classification. *Procedia engineering*, *29*, 1528–1532.

Zhu, J., Yang, Y., Xie, Q., Wang, L., & Hassan, S. U. (2014). Robust hybrid name disambiguation framework for large databases. *Scientometrics*, *98*(3), 2255–2274.

Zimmer, W. K., & Keiper, P. (2021). Redesigning curriculum at the higher education level: Challenges and successes within a sport management program. *Education Action Research*, *29*(2), 276–291.