

Received 6 November 2022, accepted 18 November 2022, date of publication 21 November 2022,  
date of current version 30 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3224034

## RESEARCH ARTICLE

# Covid-19 and E-Learning: An Exploratory Analysis of Research Topics and Interests in E-Learning During the Pandemic

FATIH GURCAN<sup>1</sup>, GONCA GOKCE MENEKSE DALVEREN<sup>2</sup>, AND MOHAMMAD DERAWI<sup>3</sup>

<sup>1</sup>Department of Computer Engineering, Faculty of Engineering, Karadeniz Technical University, 61080 Trabzon, Turkey

<sup>2</sup>Department of Software Engineering, Faculty of Engineering, Atilim University, 06830 Ankara, Turkey

<sup>3</sup>Department of Electronic Systems, Norwegian University of Science and Technology, 2815 Gjøvik, Norway

Corresponding author: Mohammad Derawi (mohammad.derawi@ntnu.no)

**ABSTRACT** E-learning has gained further importance and the amount of e-learning research and applications has increased exponentially during the COVID-19 pandemic. Therefore, it is critical to examine trends and interests in e-learning research and applications during the pandemic period. This paper aims to identify trends and research interests in e-learning articles related to COVID-19 pandemic. Consistent with this aim, a semantic content analysis was conducted on 3562 peer-reviewed journal articles published since the beginning of the COVID-19 pandemic, using the N-gram model and Latent Dirichlet Allocation (LDA) topic modeling approach. Findings of the study revealed the high-frequency bigrams such as “online learn”, “online education”, “online teach” and “distance learn”, as well as trigrams such as “higher education institution”, “emergency remote teach”, “education online learn” and “online teach learn”. Moreover, the LDA topic modeling analysis revealed 42 topics. The topics of “Learning Needs”, “Higher Education” and “Social Impact” respectively were the most focused topics. These topics also revealed concepts, dimensions, methods, tools, technologies, applications, measurement and evaluation models, which are the focal points of e-learning field during the pandemic. The findings of the study are expected to provide insights to researchers and future studies.

**INDEX TERMS** E-learning trends, topic modeling, n-grams, covid-19, topic discovery.

## I. INTRODUCTION

E-learning environments have pioneered important changes in the field of education with the opportunities and advantages they provide. E-learning specifically focuses on the online dimension of learning-teaching processes [1], [2] and is used synonymously with concepts such as online learning, virtual learning, and web-based learning [3], [4]. E-learning offers different alternatives to traditional classroom education. Namely, it allows to interact with all stakeholders synchronously and asynchronously, and allows students to securely access all course materials without time and place restrictions [5], [6]. Thanks to these important opportunities, e-learning provides educational communities with an effective learning-teaching experience, and thus it is widely accepted in all educational environments [7], [8]. For this reason, the number of articles in the field of

e-learning in the research literature is increasing exponentially every day [9].

With the COVID-19 pandemic and quarantines, e-learning has played an important role in our lives, so e-learning research during the pandemic has been critical [10], [11]. The World Health Organization (WHO) declared COVID-19 a global public health emergency on 30 January 2020 and a pandemic on 11 March 2020 [12]. COVID-19 has caused significant change and transformation in the field of education as in almost every field [13], [14], [15], [16]. In the spring of 2020, most educational institutions around the world had to suspend face-to-face education until further notice [17], [18], [19]. Thus, the face-to-face courses that form the basis of the education system have rapidly transformed into distance and online courses [18], [19], [20], [21], [22]. Thanks to Information and Communication Technologies (ICT), educational institutions in most countries have rapidly developed and implemented alternative distribution channels to move traditional education classes of almost all levels to distance

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

learning environments [10], [11], [14], [23], [24]. In this period, which is also called emergency remote teaching [25], various ICTs and e-learning opportunities have been mobilized for these e-learning applications that quickly activate distance education. In this pandemic period, all applications and activities related to distance education are characterized under the umbrella of e-learning [26].

The widespread orientation to these sub-fields of e-learning, called emergency remote teaching or emergency e-learning during the COVID-19 pandemic, has greatly increased the number of e-learning articles on the pandemic during this period [25], [27], [28]. In general terms, it is seen that the applications and activities of emergency remote teaching and emergency e-learning, which were quickly implemented during the pandemic, and the analyzes and evaluations of their different dimensions are primarily discussed in the articles. In these studies, the evaluation of the level of student satisfaction and the quality of e-learning related to e-learning activities and applications [29], [30], its effect on academic achievement [18], interest and readiness regarding e-learning [17], perception of e-learning [16], [31], e-learning adoption [32], user experiences and expectations regarding e-learning and its dimensions [20], [21], [23], [25], benefits, advantages and challenges of e-learning [21], [24], strategies for emergency e-learning and conceptual modeling [11], [15] are discussed in priority during the pandemic.

Due to the distance education requirements mandated by the current pandemic and quarantine conditions, the e-learning field has been an important focus of scientific research and practice in the pandemic. As a result, the number of domain-specific articles has increased exponentially recently. Despite this increase in the number of articles during the pandemic, there is a lack of studies that address the e-learning researches related to the pandemic from a holistic perspective with an in-depth content analysis and reveal the research landscape of the field in detail. Before the COVID-19 pandemic, various systematic reviews [2], [33], bibliometric analysis [34], [35], and content analysis studies based on topic modeling [6] were conducted, which dealt with e-learning research. These systematic reviews can only be conducted with a limited number of articles, as they are generally based on manual methods and require a lot of researcher effort [9]. Bibliometric studies can only reveal the scientific metrics and intellectual structure of any field by considering only certain bibliometric indicators such as author, article, journal, subject areas, country, and affiliation [2], but unfortunately cannot provide in-depth insights based on semantic content analysis of the domain-specific literature [9].

In this regard, analyzes in which text/data mining methods based on machine learning are used effectively on data sets containing many domain-specific articles are required for the specific determination of themes, trends and dynamics in a scientific field [6]. Topic modeling, which is a probabilistic and generative model, is the leading method used in such analyzes [36]. Unlike systematic reviews and bibliometric

analyses, the topic modeling approach enables the automatic analysis of a large scientific corpus containing a large number of articles with a systematic methodology [36]. Content analysis based on topic modeling reveals themes and trends in any scientific field from a comprehensive perspective. Although a number of topic modeling studies were carried out in the field of e-learning before the pandemic, to the best of our knowledge, a study based on topic modeling that reveals all the themes, trends and perspectives in this field with a holistic approach has not yet been carried out during the pandemic period. This current situation creates a gap in terms of in-depth analysis of the pandemic dimension in the rich literature on e-learning, which has an interdisciplinary background.

Therefore, this study aimed to bridge this gap by comprehensively revealing research interests, themes, and trends in e-learning during the pandemic. In this study, firstly, the bibliometric characteristics of the articles about e-learning in the COVID-19 period were investigated. Secondly, N-gram based content analysis was implemented, and this way, prominent keywords and terminologies in the corpus were identified. Then, an exploratory analysis was conducted using the LDA-based topic modeling approach, thus 42 trending topics were discovered that mapped the research landscape of the field. Within this context, the methodology of the study was designed to investigate the following research questions (RQ):

RQ1: What are the bibliometric characteristics of the articles?

RQ2: What are the distribution of prominent keywords and terminologies in the articles?

RQ3: What are the most focused research topics and interests in the articles?

## II. METHOD

A semi-automated methodology was proposed in this study to analyze the empirical corpus of e-learning articles during the COVID-19 pandemic. Our methodology was based on the implementation of Latent Dirichlet Allocation (LDA), a probabilistic approach for topic modeling used to discover hidden semantic structures in a domain-specific corpus. From this perspective, the research methodology designed in accordance with the study's purpose included the following stages. The empirical corpus for this study was initially prepared. Following that, data preprocessing was applied to the corpus. The LDA-based topic modeling approach was then implemented to the corpus. Finally, the topics discovered were interpreted and named. Figure 1 depicts the methodology in detail.

### A. SEARCH STRATEGY AND DATA COLLECTION

The first stage of the methodology of this study included the creation of the search strategy and the acquisition of the data constituting the experimental corpus. In this context, a search strategy was designed to include e-learning articles throughout the pandemic in the literature in order to obtain

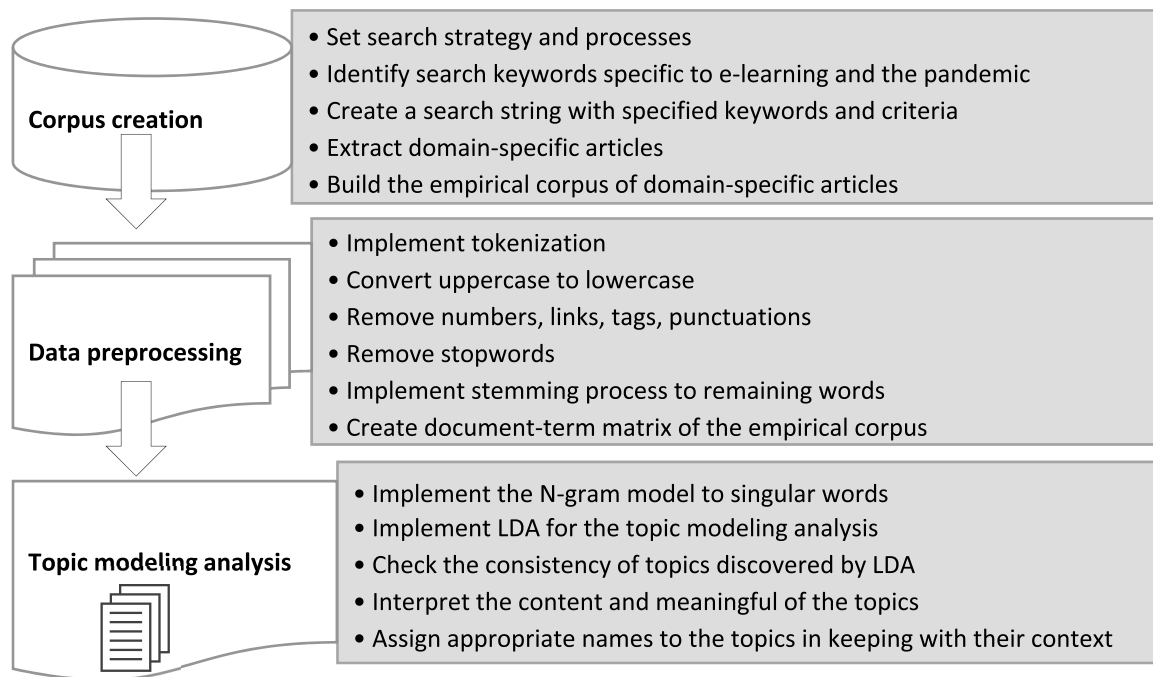


FIGURE 1. An overview of our research methodology.

the most compatible data with the background and purpose of the study. Initially, e-learning and its substitute synonyms (such as distance learning, online learning, online teaching) were determined by considering the domain-specific literature [6], [9], and these keywords were added to the search-query. These selected keywords were examined by five field experts and researchers, and the final decision was made. Then, keywords related to the pandemic (such as covid\*, pandemic\*, coronavirus\*) were added to the query. At the stage of choosing the types of publications to be included in the experimental data set of the study, only peer-reviewed journal articles were included to the data set taking into account that the journal articles are studies that have passed a certain peer-review and have reached a certain scientific maturity [37]. In addition, other criteria such as publication dates (January 2020 - June 2021) and publication language (English) were added to the query. Then, the data bibliometric database from which the articles will be obtained was chosen. The articles in this context were searched in Web of Science (WoS) and Scopus, which are considered two main data sources for bibliometric databases. As a result of search on these two sources, more articles were obtained from Scopus compared to WoS. Scopus has a broader scope of journals and includes all journals indexed by WoS [37], [38]. For this reason, the Scopus bibliometric database was chosen as the data source and the search-query was built accordingly as follows:

TITLE-ABS((covid\* OR pandemi\* OR coronavirus\*) AND (“online learn\*” OR “e-learn\*” OR “distance learn\*” OR “mobile learn\*” OR “web-based learn\*” OR “remote learn\*” OR

”virtual learn\*” OR ”digital learn\*” OR “blend\* learn\*” OR “online train\*” OR “e-train\*” OR “distance train\*” OR “mobile train\*” OR “web-based train\*” OR ”remote train\*” OR ”virtual train\*” OR ”digital train\*” OR ”blend\* train\*” OR ”online educat\*” OR “e-educat\*” OR “distance educat\*” OR “mobile educat\*” OR “web-based educat\*” OR “remote educat\*” OR “virtual educat\*” OR ”digital educat\*” OR ”blend\* educat\*” OR “online teach\*” OR “e-teach\*” OR “distance teach\*” OR “mobile teach\*” OR “web-based teach\*” OR “remote teach\*” OR ”virtual teach\*” OR ”digital teach\*” OR “blend\* teach\*” OR “MOOC” OR “online open course” OR “online course” OR ”online classroom\*” OR ”virtual classroom\*”)) AND (LIMIT-TO (DOCTYPE, “ar”) OR LIMIT-TO (DOCTYPE, “re”)) AND (LIMIT-TO (LANGUAGE, “English”))

Using this search-query, a total of 3562 articles were obtained as a result of the search in the Scopus database on July 6, 2021 (3371 research articles, 191 reviews). Then, an experimental corpus was created, consisting of the title, abstract, author keywords and publication year of these articles. Because the title, abstract, and author keywords of an article are the most important sections that definitely reveal the background, scope, method and purpose of an article [39].

**B. DATA PREPROCESSING**

Data preprocessing is a critical task that directly affects the achievement of the analysis, especially for text mining-based studies [40], [41]. For this reason, some

preprocessing operations required for this corpus were implemented sequentially. Initially, all texts were converted to lowercase. Then, web links, publisher information, numeric expressions, punctuation, and symbols were deleted. Word tokenization was applied. In the next step, the English stop words (the, are, and, is, or, a, an, for, etc.) were deleted. In order to clear the words in the corpus from their derived forms and reduce them to the root, stemming process was implemented using the Snowball rooting algorithm [42]. Finally, each article in the corpus was modeled as a word vector based on the “bag of words” model to provide a numerical representation of the corpus. The article texts represented by these vectors were transformed into a document term matrix (DTM). Thus, the corpus was converted into a suitable numerical matrix form for the topic modeling analysis [36], [41].

In addition, content analysis based on N-gram model for the word-level was performed in order to reveal the domain-specific semantic structures and terminologies that are frequently mentioned in the corpus [9]. In the N-gram model, unigrams represent single words, bigrams represent sequences of two words, trigrams represent sequences of three words, and so on. Using this approach, the frequencies of unigrams, bigrams, and trigrams in the corpus were calculated, accordingly prominent semantic structures and terminologies in the corpus were revealed.

### C. FITTING AND IMPLEMENTING THE TOPIC MODELING

Topic modeling is a machine learning approach often used in text mining and natural language processing to discover hidden semantic structures known as “topics” in a body of text [36]. The background of topic modeling is based on probabilistic approaches and statistical algorithms. LDA (Latent Dirichlet Allocation) algorithm, which uses probabilistic and generative processes methodologically, is an effective topic modeling technique for semantic content analysis [43]. LDA is a machine learning algorithm that discovers hidden topics from a text with an approach based on unsupervised learning and does not require any training process, thus enabling systematic computational analysis of large numbers of text documents [36], [38], [43]. In this study, the MALLET package [44], which provides a basic infrastructure based on Gibbs sampling iteration, was used to fit and implement LDA topic modeling to the experimental corpus. The LDA model was implemented with various K topic number values, and consistent topics with optimal topic-word distributions were obtained when the number of topics K was equal to 42. Each of these topics was represented by the top 15 descriptive keywords that reflect the scope of the topics.

### III. FINDINGS

The results of the study are firstly introduced descriptively in order to show bibliometric characteristics. Further, N-gram and topic modeling analyses are presented to provide an overall portrait of e-learning studies in the period of COVID-19 pandemic.

### A. DESCRIPTIVE ANALYSIS

In order to answer the Research Question 1 (RQ1) a total of 3562 articles were analyzed (191 of them are literature reviews and 3371 of them are research articles). 1381 of the articles were published in 2020 while 2181 of them were published in 2021; and their distribution by the subject areas are given in Table 1. Since some of the articles are indexed in more than one subject area, the total number of articles by the subject areas and their total rate should not be a misleading issue.

**TABLE 1. Distribution of E-learning studies published during COVID-19 by their subject areas.**

| Subject Areas                                | N    | %     |
|--|------|-------|
| Social Sciences                              | 2132 | 59.85 |
| Medicine                                     | 863  | 24.23 |
| Computer Science                             | 536  | 15.05 |
| Psychology                                   | 291  | 8.17  |
| Engineering                                  | 241  | 6.77  |
| Arts and Humanities                          | 233  | 6.54  |
| Environmental Science                        | 216  | 6.06  |
| Business, Management and Accounting          | 162  | 4.55  |
| Biochemistry, Genetics and Molecular Biology | 152  | 4.27  |
| Chemistry                                    | 149  | 4.18  |
| Health Professions                           | 146  | 4.10  |
| Nursing                                      | 127  | 3.57  |
| Energy                                       | 112  | 3.14  |
| Agricultural and Biological Sciences         | 97   | 2.72  |
| Pharmacology, Toxicology and Pharmaceutics   | 92   | 2.58  |
| Mathematics                                  | 76   | 2.13  |
| Economics, Econometrics and Finance          | 69   | 1.94  |
| Dentistry                                    | 65   | 1.82  |
| Physics and Astronomy                        | 62   | 1.74  |
| Multidisciplinary                            | 50   | 1.40  |
| Immunology and Microbiology                  | 46   | 1.29  |
| Materials Science                            | 39   | 1.09  |
| Neuroscience                                 | 37   | 1.04  |
| Chemical Engineering                         | 34   | 0.95  |
| Decision Sciences                            | 30   | 0.84  |
| Earth and Planetary Sciences                 | 16   | 0.45  |
| Veterinary                                   | 12   | 0.34  |

When Table 1 is examined, it is observed that “Social Sciences”, “Medicine” and “Computer Science” are the most published subject areas, while “Earth and Planetary Sciences” and “Veterinary” are the least published. The distribution of the articles by the top 20 journals is given in Table 2.

Examined in detailed, Table 2 reveals that the journals with the highest number of articles published are “Journal of Chemical Education”, “Sustainability Switzerland” and “Education Sciences” respectively. The top 20 countries with the highest number of articles originated from are enlisted in Table 3. Table 3 informs that the highest number of articles originated from “United States”, “India” and “United Kingdom”, respectively.

**TABLE 2. Top 20 journals where E-learning studies published during COVID-19 period.**

| Source Titles   | N   | %    |
|---|-----|------|
| Journal of Chemical Education                                       | 120 | 3.37 |
| Sustainability Switzerland  | 100 | 2.81 |
| Education Sciences  | 62  | 1.74 |
| Frontiers In Psychology   | 50  | 1.40 |
| Education and Information Technologies                              | 49  | 1.38 |
| Gms Journal for Medical Education                                   | 44  | 1.24 |
| International Journal of Environmental Research and Public Health   | 43  | 1.21 |
| Frontiers in Education  | 33  | 0.93 |
| Journal of Microbiology and Biology Education                       | 33  | 0.93 |
| Educational Technology Research and Development                     | 31  | 0.87 |
| Information and Learning Science                                    | 28  | 0.79 |
| Biochemistry and Molecular Biology Education                        | 27  | 0.76 |
| BMC Medical Education   | 25  | 0.70 |
| International Journal of Learning Teaching and Educational Research | 24  | 0.67 |
| Medical Science Educator  | 23  | 0.65 |
| Plos One  | 23  | 0.65 |
| Ecology and Evolution   | 21  | 0.59 |
| Journal of Engineering Education Transformations                    | 21  | 0.59 |
| The Journal of National Education                                   | 21  | 0.59 |
| Communications of the Association for Information Systems           | 20  | 0.56 |

**TABLE 3. Distribution of E-learning studies published during COVID-19 by the top 20 countries.**

| Countries          | N   | %     |
|--------------------|-----|-------|
| United States      | 992 | 27.85 |
| India              | 240 | 6.74  |
| United Kingdom     | 220 | 6.18  |
| China              | 192 | 5.39  |
| Saudi Arabia       | 162 | 4.55  |
| Australia          | 153 | 4.30  |
| Indonesia          | 150 | 4.21  |
| Canada             | 141 | 3.96  |
| Spain              | 137 | 3.85  |
| Italy              | 116 | 3.26  |
| Germany            | 111 | 3.12  |
| Malaysia           | 109 | 3.06  |
| Turkey             | 91  | 2.55  |
| Pakistan           | 72  | 2.02  |
| South Africa       | 72  | 2.02  |
| Brazil             | 65  | 1.82  |
| Russian Federation | 63  | 1.77  |
| Jordan             | 61  | 1.71  |
| Poland             | 61  | 1.71  |
| Hong Kong          | 56  | 1.57  |

**B. N-GRAM BASED CONTENT ANALYSIS**

This section includes the result of N-gram (unigram, bigram and trigram) analysis employed to identify the high frequency terms in the corpus of e-learning studies during COVID-19.

With regard to the Research Question 2 (RQ2), the results of unigram, bigram, and trigram analyses are given in Table 4, Table 5, and Table 6.

**TABLE 4. Top 20 ranked unigrams in the articles.**

| Unigrams    | N    | %     |
|-------------|------|-------|
| learn       | 2909 | 81.67 |
| online      | 2539 | 71.28 |
| education   | 2520 | 70.75 |
| student     | 2429 | 68.19 |
| teach       | 1840 | 51.66 |
| experience  | 1298 | 36.44 |
| university  | 1296 | 36.38 |
| course      | 1180 | 33.13 |
| educational | 1134 | 31.84 |
| data        | 1033 | 29.00 |
| method      | 967  | 27.15 |
| impact      | 966  | 27.12 |
| technology  | 958  | 26.90 |
| survey      | 950  | 26.67 |
| school      | 936  | 26.28 |
| teacher     | 934  | 26.22 |
| distance    | 917  | 25.74 |
| remote      | 903  | 25.35 |
| support     | 896  | 25.15 |
| social      | 876  | 24.59 |

As seen in Table 4, the highest ranked unigram in the corpus is the term “learn” with the rate of 81.67%. In other words, a total of 2909 documents have included the term “learn”. It is followed by the term “online” (n=2539; f=71.28%) and “education” (n=2520, f=70.75%).

**TABLE 5. Top 20 ranked bigrams in the articles.**

| Bigrams                 | %    | N     |
|-------------------------|------|-------|
| online learn            | 1577 | 44.27 |
| online education        | 722  | 20.27 |
| online teach            | 652  | 18.30 |
| distance learn          | 635  | 17.83 |
| higher education        | 613  | 17.21 |
| student learn           | 506  | 14.21 |
| remote learn            | 383  | 10.75 |
| educational institution | 370  | 10.39 |
| social distance         | 518  | 14.54 |
| medical education       | 308  | 8.65  |
| learn environment       | 301  | 8.45  |
| remote teach            | 275  | 7.72  |
| online course           | 261  | 7.33  |
| distance education      | 258  | 7.24  |
| learn experience        | 232  | 6.51  |
| data collect            | 218  | 6.12  |
| virtual learn           | 217  | 6.09  |
| university student      | 212  | 5.95  |
| medical student         | 207  | 5.81  |
| learn process           | 206  | 5.78  |

As shown in Table 5, the highest ranked bigram in the corpus is “online learn” two-word sequence with the rate of 44.27%. Specifically, the two-word sequence of “online learn” have occurred in a total of 1577 documents. It is followed by “online education” (n=722; f=20.27%) and “online teach” (n=652, f=18.30%) two-word sequences.

**TABLE 6. The top ranked 20 trigrams in the articles.**

| Trigrams                             | N   | %    |
|--------------------------------------|-----|------|
| higher education institution         | 151 | 4.24 |
| emergency remote teach               | 137 | 3.85 |
| education online learn               | 119 | 3.34 |
| online teach learn                   | 117 | 3.28 |
| division chemical education          | 106 | 2.98 |
| online learn environment             | 85  | 2.39 |
| learn management system              | 81  | 2.27 |
| higher education online              | 81  | 2.27 |
| learn self instruction               | 76  | 2.13 |
| information communication technology | 76  | 2.13 |
| learn higher education               | 76  | 2.13 |
| online learn student                 | 73  | 2.05 |
| student online learn                 | 68  | 1.91 |
| transition online learn              | 57  | 1.60 |
| online learn experience              | 55  | 1.54 |
| distance learn self                  | 50  | 1.40 |
| internet web-based learn             | 49  | 1.38 |
| virtual learn environment            | 43  | 1.21 |
| massive online course                | 41  | 1.15 |
| transition remote learn              | 39  | 1.09 |

As seen in the Table 6, the highest ranked trigram in the corpus is the three-word sequence of “higher education institution” with the rate of 4.24%. Specifically, this three-word sequence has occurred in 151 documents. It is followed by three-word sequences of “emergency remote teach” (n=137; f=3.85%) and “education online learn” (n=119, f=3.34%).

### C. TOPIC MODELING ANALYSIS

Finally, regarding the Research Question 3 (RQ3) the results of topic modeling analysis with LDA are given in Table 7. As a result, 42 main topics are discovered. The rate of each topic is calculated considering the number of the articles in that topic. Besides, the top 15 keywords for each topic are given in the table.

When Table 7 is analyzed in detail, it is observed that during the period of COVID-19 the most studied topic is “Learning Needs” with the rate of 6.7%. It is followed by “Higher education” and “Social impact” with the rates of 5.94% and 4.61%, respectively. It is noticed that the least studied topic is “Disability training” with the rate of 0.55%.

## IV. DISCUSSION

This study aimed to identify the current trends and research interests in e-learning studies related to COVID-19 pandemic. In this respect, bibliometric characteristics of 3562 articles were firstly determined; then high frequency terms and topics

in the articles were revealed using N-gram and LDA-based topic modeling approaches. Considering the number of published articles in 2020 and 2021 (almost midyear) as 1381 and 2181, respectively. As highlighted in previous studies, distance education and e-learning have quickly become popular during the COVID-19 pandemic [15], [21], [25], [32] and the number of articles in this field have accelerated in recent years [34], [35]. Based on the subject areas of the articles, it can be concluded that there have been publications almost in every field. Namely, “Social Sciences”, “Medicine” and “Computer Science” are the top ranked subject areas with highest number of articles. Having in mind educational sciences is classified under social sciences [9], [35], the highest ranks of “Social Sciences” and “Medicine” support the previous findings [35], [45]. Regarding the distribution of the articles by the journals, among a wide range of journals the top ranked ones with the highest number of articles published are revealed as “Journal of Chemical Education”, “Sustainability Switzerland” and “Education Sciences”. Although many previous studies [35] have concluded that articles in the field of e-learning are mostly published in the “Computers and Education” journal, this study reveals the different journals. Considering the origins of the articles, “United States”, “India” and “United Kingdom” are ranked as top three, respectively. In accordance with the previous studies [9], [34] this study also finds out “United States” as the top ranked country with the highest number of articles originated. Besides, countries such as “United States”, “United Kingdom” and “China” are generally top ranked countries in the literature [9] and in this study as well.

N-gram analyses (unigram, bigram, and trigram) were employed in order to determine the frequent terms in the articles. With these analyses, it was aimed to identify prominent keywords and high frequency terms in the corpus and to determine domain specific contexts. The results of unigrams reveal the terms such as “learn”, “online”, “education” which are directly related to e-learning. Further examination of unigrams shows us that there are also keywords related to e-learning and distance learning like “online learn”, “online education”, “online teach”, “distance learn”. The trigrams such as “higher education institution”, “emergency remote teach”, “education online learn” and “online teach learn” are also frequently highlighted in the corpus. The first trigram, “higher education institution”, is an important indicator of the acceleration of e-learning studies in higher education during the pandemic [15], [23], [24], [32]. The remaining trigrams can be evaluated as directly indicators of the rapid transition to e-learning and distance education [20], [25], and so the results of the current study support this.

Finally, topic modeling analysis of e-learning articles published during COVID-19 revealed 42 main topics. It was observed that “Learning Needs”, “Higher Education”, “Social Impact” and “Blended Learning” are the leading topics. These discovered topics also revealed the concepts and methods of e-learning (e.g. topics “Online Teaching”, “Distance Education”, “Emergency Remote

**TABLE 7. Discovered topics with their percentages and related keywords.**

| Topic Name                | Descriptive Keywords  | Rate (%) |
|---------------------------|---|----------|
| Learning Needs            | learn need development develop skill resource program strategy design learner                           | 6.57     |
| Higher Education          | education higher challenge online institution university teach educational learn technology             | 5.94     |
| Social Impact             | impact social change life affect crisis effect time education world                                     | 4.61     |
| Blended Learning          | learn blend online learner environment experience perception face-to-face explore shift                 | 4.54     |
| Practice Learning         | practice learn education teach community online educator design pedagogy value                          | 4.10     |
| Online Courses            | course online learn instructor transition instruction teach synchronous experience design               | 4.06     |
| Online Teaching           | online teach class classroom course traditional college mode tool effective                             | 3.80     |
| Survey                    | survey response participant lecture online respondent report education asses questionnaire              | 3.57     |
| Educational Development   | education university educational process teach development higher sustainable situation analysis        | 3.05     |
| Distance Education        | distance education learn process form activity period teacher opinion problem                           | 3.02     |
| Qualitative Data Analysis | datum analysis qualitative interview experience participant collect understand analyze theme            | 2.89     |
| Readiness Level Analysis  | analysis level readiness attitude perception datum online questionnaire factor sample                   | 2.84     |
| Mental Health             | anxiety mental health stress psychological impact level cope depression university                      | 2.76     |
| Emergency Remote Teaching | remote emergency learn teach experience educational transition ert response crisis                      | 2.60     |
| Teacher Education         | teacher teach education online school mathematic professional technology development practice           | 2.47     |
| E-Learning                | e-learn university learn system internet lecturer management barrier lms challenge                      | 2.44     |
| Public Health             | health disease public measure infection risk emergency personal control equipment                       | 2.27     |
| Chemistry Education       | laboratory chemistry learn chemical undergraduate lab instruction american general society              | 2.23     |
| Satisfaction Factors      | factor satisfaction perceive self-efficacy quality intention acceptance influence technology education  | 2.18     |
| Resident Education        | resident education program surgical trainee train virtual surgery medical residency                     | 2.14     |
| Virtual Classroom         | virtual learn classroom session conference video environment zoom experience group                      | 2.12     |
| School Learning           | school learn closure secondary teacher elementary primary high close district                           | 2.05     |
| Online Faculties          | faculty university member online teach college staff program transition experience                      | 2.03     |
| Medical Education         | medical education dental clinical school teach medicine undergraduate curriculum clerkship              | 1.99     |
| Student Engagement        | engagement online learn motivation social cognitive presence support emotional positive                 | 1.94     |
| ICT                       | digital technology ict competence education information communication tool technological transformation | 1.85     |
| Mobile Learning           | system mobile application reality datum learn technology device network solution                        | 1.78     |
| Access Policy             | access policy inequality rural education country global equity international gap                        | 1.71     |
| Performance Assessment    | assessment performance academic exam test online examination grade feedback university                  | 1.62     |
| Patient-Care Training     | train care patient train clinical service practice therapy support telehealth                           | 1.51     |
| Physical Activity         | physical activity sleep time adolescent adult day sport change screen                                   | 1.40     |
| Social Media              | social media search information post network tool facebook internet web                                 | 1.40     |
| Team Management           | management team crisis business job role service change industry system                                 | 1.39     |
| Child Education           | child parent family school children support home adolescent childhood online                            | 1.30     |
| EFL                       | language english efl foreign write teach learner university teacher read                                | 1.27     |
| Experimental Study        | group intervention control test score participant train think knowledge experimental                    | 1.19     |
| Digital Medical Teaching  | teach medical digital clinical format train patient concept course skill                                | 1.09     |
| Library Services          | design library university hybrid service information project academic support literacy                  | 0.98     |
| MOOC                      | mooc course video massive podcast content topic youtube analytic user                                   | 0.97     |
| Nursing Education         | nursing nurse education simulation health clinical care interprofessional practice ne                   | 0.91     |
| Gamification              | game gamification lecture simulation engineer physic application serious robotic concept                | 0.89     |
| Disability Training       | disability special art people visual train dance ne challenge live                                      | 0.55     |

Teaching”, “Mobile Learning”, and “Blended Learning”), tools and technologies used in e-learning environments (e.g. topics “Virtual Classroom”, “Gamification”, “Online Courses”, “MOOC”, “ICT”, “Library Services”, “Social Media”), educational areas where e-learning is applied (e.g. topics “Higher Education”, “Teacher Education”, “Chemistry Education”, “Resident Education”, “Online Faculties”, “Medical Education”, “Child Education”, “Nursing Education”, “EFL”, “Digital Medical Teaching”, “Patient-

Care Training”, and “Disability Training”), measurement and evaluation models used in e-learning environments (e.g. topics “Qualitative Data Analysis”, “Readiness Level Analysis”, “Satisfaction Factors”, “Survey”, “Experimental Study”, “Student Engagement”, “Performance Assessment”, and “Educational Development”), and health dimension of e-learning (e.g. topics “Mental Health”, “Physical Activity”, “Public Health”). In this way, the discovered topics provide important insights into understanding

the most focused motivations, applications, and dimensions in the field [14], [18], [24].

## V. CONCLUSION AND FUTURE STUDIES

In this study, considering the e-learning studies related to the COVID-19 pandemic, it was aimed to reveal the research landscape described by these studies. In the study, which is built on three pillars, initially, descriptive analyzes were performed to identify the bibliometric features of the field. After that, by using N-gram analyses, the most frequent terms in the context were identified to determine domain specific context. Finally, by implementing LDA-based topic modeling analysis, topic distributions of the articles were analyzed. The results of the study are significant as they present the existing situation and research trends in e-learning field during COVID-19. The findings on bibliometric characteristics of the field demonstrate the followings: e-learning studies accelerated during COVID-19; the journals where the articles are published are considerably diverse; the studies are conducted in different fields and disciplines, particularly social sciences; “United States” is the first ranked country in terms of origin. The results of N-gram analyses reflect the main frame of distance learning and e-learning. Trigram results particularly unfold that emergency distance education and studies in higher education are prominent. Finally, the results of topic modeling analysis reveal that topics such as “Learning Needs”, “Higher Education”, “Social Impact” and “Blended Learning” are prominent ones.

Our findings show that the educational precautions in the early stages of COVID-19 focus on using distance education strategies and procedures as an emergency response. As the pandemic progressed, so did distance education activities. A complete understanding of the pandemic’s short, medium, and long-term effects is still developing. Our study explored how the pandemic has impacted distance learning and traditional learning environments. Our findings revealed the e-learning concepts, methods, tools and technologies used in e-learning environments during the pandemic. The findings also demonstrated common e-learning activities, application areas, measurement and evaluation models used in e-learning, and social and health dimensions of e-learning, all of which are focal points of the e-learning field during the pandemic.

In this study, descriptive analysis, N-gram-based content analysis, and LDA-based topic modeling analysis were performed on a total of 3562 articles published during COVID-19. Considering that COVID-19 pandemic continues, the results of the current study are expected to provide a basis and insights for future studies. An automated text mining methodology based on generative topic modeling proposed in this study can be applied to different sub-contexts of educational sciences to obtain up-to-date implications for practitioners and educators in various backgrounds. Educational regulations and policies can be updated by taking into account the trends and themes in e-learning field revealed in this study. In today’s e-learning discipline, where the

number of publications is rapidly increasing, such content analysis based on unsupervised machine learning may be a more useful methodology and guide for the ever-expanding education communities in the near future. Taking these trends in emergency e-learning strategies into account, training programs for in-service teachers and educators can be updated, thus providing candidates with emergency response skills for pandemics. Furthermore, the results of the study can lead to detailed investigations through performing more comprehensive analysis on particular topics.

## REFERENCES

- [1] N. Jayakumar, O. Brunckhorst, P. Dasgupta, M. S. Khan, and K. Ahmed, “E-learning in surgical education: A systematic review,” *J. Surgical Educ.*, vol. 72, no. 6, pp. 1145–1157, Nov. 2015.
- [2] J. Valverde-Berrococo, M. D. C. Garrido-Arroyo, C. Burgos-Videla, and M. B. Morales-Cevallos, “Trends in educational research about E-learning: A systematic literature review (2009–2018),” *Sustainability*, vol. 12, no. 12, p. 5153, Jun. 2020.
- [3] A. Chiappe and L. L. Lee, “Open teaching: A new way on E-learning?” *Electron. J. E-Learn.*, vol. 15, no. 5, pp. 369–383, 2017.
- [4] E. O’Donnell, S. Lawless, M. Sharp, and V. P. Wade, “A review of personalised E-learning: Towards supporting learner diversity,” *Int. J. Distance Educ. Technol.*, vol. 13, no. 1, pp. 22–47, 2015.
- [5] M. Derawi, “E-learning protection of open access platforms,” in *Proc. Int. Conf. Web Open Access Learn. (ICWOAL)*, Nov. 2014, pp. 1–6.
- [6] F. Gurcan, O. Ozyurt, and N. E. Cagiltay, “Investigation of emerging trends in the E-learning field using latent Dirichlet allocation,” *Int. Rev. Res. Open Distrib. Learn.*, vol. 22, no. 2, pp. 1–18, Jan. 2021.
- [7] H. Al-Samarraie, H. Selim, T. Teo, and F. Zaquout, “Isolation and distinctiveness in the design of E-learning systems influence user preferences,” *Interact. Learn. Environ.*, vol. 25, no. 4, pp. 452–466, May 2017.
- [8] L. Shahmoradi, V. Changizi, E. Mehraeen, A. Bashiri, B. Jannat, and M. Hosseini, “The challenges of E-learning system: Higher educational institutions perspective,” *J. Educ. Health Promotion*, vol. 7, p. 116, Jan. 2018.
- [9] F. Gurcan and N. E. Cagiltay, “Research trends on distance learning: A text mining-based literature review from 2008 to 2018,” *Interact. Learn. Environ.*, vol. 2020, pp. 1–22, Sep. 2020.
- [10] M. A. Almaiah, A. Al-Khasawneh, and A. Althunibat, “Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic,” *Educ. Inf. Technol.*, vol. 25, pp. 5261–5280, Nov. 2020.
- [11] M. A. Adarkwah, “I’m not against online teaching, but what about us?: ICT in Ghana post COVID-19,” *Educ. Inf. Technol.*, vol. 26, no. 2, pp. 1665–1685, Mar. 2021.
- [12] *Addressing Human Rights as Key to the COVID-19 Response*. World Health Organization, Geneva, Switzerland, 2020.
- [13] G. Kaisara and K. J. Bwalya, “Investigating the E-learning challenges faced by students during COVID-19 in Namibia,” *Int. J. Higher Educ.*, vol. 10, no. 1, p. 308, Nov. 2020.
- [14] P. Paudel, “Online education: Benefits, challenges and strategies during and after COVID-19 in higher education,” *Int. J. Stud. Educ.*, vol. 3, no. 2, pp. 70–85, Sep. 2020.
- [15] A. Shahzad, R. Hassan, A. Y. Aremu, A. Hussain, and R. N. Lodhi, “Effects of COVID-19 in E-learning on higher education institution students: The group comparison between male and female,” *Quality Quantity*, vol. 55, no. 3, pp. 805–826, Jun. 2021.
- [16] A. Kamble, R. Gauba, S. Desai, and D. Golhar, “Learners’ perception of the transition to instructor-led online learning environments: Facilitators and barriers during the COVID-19 pandemic,” *Int. Rev. Res. Open Distrib. Learn.*, vol. 22, no. 1, pp. 199–215, Mar. 2021.
- [17] A. R. Prasetyo, H. Nurtjahjanti, and L. N. Ardiani, “Impact of changes in teaching methods during the COVID-19 pandemic: The effect of integrative E-learning on readiness for change and interest in learning among Indonesian university students,” *Int. Rev. Res. Open Distrib. Learn.*, vol. 22, no. 2, pp. 87–101, Jan. 2021.
- [18] G. Ghaleb, A. E. Refae, A. Kaba, and S. Eletter, “The the impact of demographic characteristics on academic performance: Face-to-face learning versus distance learning implemented to prevent the spread of COVID-19,” *Int. Rev. Res. Open Distrib. Learn.*, vol. 22, no. 1, pp. 91–110, Mar. 2021.



- [19] M. Rizun and A. Strzelecki, "Students' acceptance of the COVID-19 impact on shifting higher education to distance learning in Poland," *Int. J. Environ. Res. Public Health*, vol. 17, no. 18, pp. 1–19, 2020.
- [20] K. Kulikowski, S. Przytula, and Ł. Sulkowski, "E-learning? Never again! On the unintended consequences of COVID-19 forced E-learning on academic teacher motivational job characteristics," *Higher Educ. Quart.*, vol. 76, no. 1, pp. 174–189, Jan. 2022.
- [21] A. M. Maatuk, E. K. Elberkawi, S. Aljawarneh, H. Rashaideh, and H. Alharbi, "The COVID-19 pandemic and E-learning: Challenges and opportunities from the perspective of students and instructors," *J. Comput. Higher Educ.*, vol. 34, no. 1, pp. 21–38, Apr. 2022.
- [22] Y. Kotera, P. Green, C. Rhodes, A. Williams, J. Chircop, R. Spink, R. Rawson, and U. Okere, "Dealing with isolation using online morning huddles for university lecturers during physical distancing by COVID-19: Field notes," *Int. Rev. Res. Open Distrib. Learn.*, vol. 21, no. 4, pp. 238–244, Jun. 2020.
- [23] K. Cicha, M. Rizun, P. Rutecka, and A. Strzelecki, "COVID-19 and higher education: First-year students' expectations toward distance learning," *Sustainability*, vol. 13, no. 4, pp. 1–20, 2021.
- [24] D. Turnbull, R. Chugh, and J. Luck, "Transitioning to E-learning during the COVID-19 pandemic: How have higher education institutions responded to the challenge?" *Educ. Inf. Technol.*, vol. 26, no. 5, pp. 6401–6419, Sep. 2021.
- [25] A. M. Müller, C. Goh, L. Z. Lim, and X. Gao, "COVID-19 emergency elearning and beyond: Experiences and perspectives of university educators," *Educ. Sci.*, vol. 11, no. 1, pp. 1–15, 2021.
- [26] V. Singh and A. Thurman, "How many ways can we define online learning? A systematic literature review of definitions of online learning (1988–2018)," *Amer. J. Distance Educ.*, vol. 33, no. 4, pp. 289–306, Oct. 2019.
- [27] A. Y. Alqahtani and A. A. Rajkhan, "E-learning critical success factors during the COVID-19 pandemic: A comprehensive analysis of E-learning managerial perspectives," *Educ. Sci.*, vol. 10, no. 9, pp. 1–16, 2020.
- [28] S. Dhawan, "Online learning: A panacea in the time of COVID-19 crisis," *J. Educ. Technol. Syst.*, vol. 49, no. 1, pp. 5–22, Sep. 2020.
- [29] E. Aboagye, J. A. Yawson, and K. N. Appiah, "COVID-19 and E-learning: The challenges of students in tertiary institutions," *Social Educ. Res.*, vol. 2020, pp. 109–115, Jun. 2020.
- [30] C. Saxena, H. Baber, and P. Kumar, "Examining the moderating effect of perceived benefits of maintaining social distance on E-learning quality during COVID-19 pandemic," *J. Educ. Technol. Syst.*, vol. 49, no. 4, pp. 532–554, Jun. 2021.
- [31] D. A. Hamaidi, Y. M. Arouri, R. K. Noufa, and I. T. Aldrou, "Parents' perceptions of their children's experiences with distance learning during the COVID-19 pandemic," *Int. Rev. Res. Open Distrib. Learn.*, vol. 22, no. 2, pp. 224–241, 2021.
- [32] M. Mailizar, D. Burg, and S. Maulina, "Examining university students' behavioural intention to use E-learning during the COVID-19 pandemic: An extended TAM model," *Educ. Inf. Technol.*, vol. 26, no. 6, pp. 7057–7077, Nov. 2021.
- [33] M. W. Rodrigues, S. Isotani, and L. E. Zárate, "Educational data mining: A review of evaluation process in the E-learning," *Telematics Informat.*, vol. 35, no. 6, pp. 1701–1717, Sep. 2018.
- [34] S. Asadzandi, T. Rakhshani, and A. Mohammadi, "Content analysis study of E-learning literature based on Scopus record through 2013: With a focus on the place of Iran's productions," *Int. J. E-Learn. Corp. Gov. Heal. High. Educ.*, vol. 16, no. 3, pp. 213–229, 2017.
- [35] G. Tibaná-Herrera, M. T. Fernández-Bajón, and F. De Moya-Anegón, "Categorization of E-learning as an emerging discipline in the world publication system: A bibliometric study in scopus," *Int. J. Educ. Technol. Higher Educ.*, vol. 15, no. 1, p. 21, Dec. 2018.
- [36] D. M. Blei, "Probabilistic topic models," *Commun. ACM*, vol. 55, no. 4, pp. 77–84, Apr. 2012.
- [37] P. Mongeon and A. Paul-Hus, "The journal coverage of web of science and Scopus: A comparative analysis," *Scientometrics*, vol. 106, no. 1, pp. 213–228, Jan. 2016.
- [38] F. Gurcan and N. E. Cagiltay, "Exploratory analysis of topic interests and their evolution in bioinformatics research using semantic text mining and probabilistic topic modeling," *IEEE Access*, vol. 10, pp. 31480–31493, 2022.
- [39] F. Gurcan, G. G. M. Dalveren, N. E. Cagiltay, and A. Soylu, "Detecting latent topics and trends in software engineering research since 1980 using probabilistic topic modeling," *IEEE Access*, vol. 10, pp. 74638–74654, 2022.
- [40] F. Gurcan, G. G. M. Dalveren, N. E. Cagiltay, D. Roman, and A. Soylu, "Evolution of software testing strategies and trends: Semantic content analysis of software research corpus of the last 40 years," *IEEE Access*, vol. 10, pp. 106093–106109, 2022.
- [41] F. Gurcan, "Extraction of core competencies for big data: Implications for competency-based engineering education," *Int. J. Eng. Educ.*, vol. 35, no. 4, pp. 1110–1115, 2019.
- [42] M. F. Porter, "Snowball: A language for stemming algorithms," Tech. Rep., 2001.
- [43] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.
- [44] A. K. McCallum. (2002). *MALLET: A Machine Learning for Language Toolkit*. [Online]. Available: <http://Mallet.Cs.Umass.Edu>
- [45] S. Barteit, D. Guzek, A. Jahn, T. Bärnighausen, M. M. Jorge, and F. Neuhann, "Evaluation of E-learning for medical education in low- and middle-income countries: A systematic review," *Comput. Educ.*, vol. 145, Feb. 2020, Art. no. 103726.



**FATIH GURCAN** received the Ph.D. degree in computer engineering from Karadeniz Technical University. He was an Instructor at the Department of Informatics, Karadeniz Technical University, from 2001 to 2014. He has been with the Center for Research and Application in Distance Education, Karadeniz Technical University, since 2015, as an Instructor. His research interests include trend analysis, sentiment analysis, statistical topic modeling, engineering education, data mining, machine learning, big data analytics, and text mining.



**GONCA GOKCE MENEKSE DALVEREN** received the Ph.D. degree in software engineering from Atilim University, and the Postdoctoral degree from the Department of Computer Science, Norwegian University of Science and Technology. Currently, she is with the Software Engineering Department, Atilim University. Her research interests include software engineering, eye tracking, medical informatics, and human-computer interaction.



**MOHAMMAD DERAWI** received the Diploma degree in computer science engineering and the B.Sc. and M.Sc. degrees from the Technical University of Denmark (DTU), Denmark, in 2007 and 2009, respectively, and the Ph.D. degree in information security from the Norwegian Information Security Laboratory (NISLab), Gjøvik University (now NTNU), Norway. His Ph.D. research interests include smart mobile technologies and also biometrics with specialization on behavioral biometric recognition in mobile devices. In 2009, he received the title as the Youngest Engineer of Denmark. In the beginning of his Ph.D. studies, he was a Visiting Researcher at the Center for Advanced Security Research Darmstadt (CASED), [www.cased.de](http://www.cased.de), Germany. He is currently having a dual-career acting as a Youngest Professor of Norway and an Extremely True Innovator. He is also the Head of the "Smart Wireless Systems" Research Group, Department of Electronic Systems, NTNU, Norway. Since 2009, he has been active in several European and national projects. Today he holds a professorship within electrical engineering. He has next to his academic career also been working with truly innovative projects which are also a part of his additional skills. He is a person that identifies the need within an industry, market segment or culture, and spot opportunity in it. More importantly, he has also the ability to identify needs before implemented in the market, develop and refine solutions, take chances, push the envelope, and create meaning. He is specialized within information security, e-health, autonomous systems, biometric systems, wireless communications, the IoT, and digital fundamentals micro controllers.