Journal of Universal Computer Science, vol. 29, no. 6 (2023), 510-545
submitted: 9/6/2022, accepted: 15/2/2023, appeared: 28/6/2023 CC BY-ND 4.0

A Bibliometric Study on E-Learning Software Engineering Education

Soukaina Benabdelouahab
(Abdelmalek Essaadi University, Tetouan, Morocco,
https://orcid.org/0000-0003-2751-6763, benabdelouahab.soukaina@gmail.com)

José A. García-Berná
(University of Murcia, Spain,
https://orcid.org/0000-0002-9526-8565, josealberto.garcia1@um.es)

Chaimae Moumouh
(Abdelmalek Essaadi University, Tetouan, Morocco,
https://orcid.org/0000-0001-9999-847X, chaimae.moumouh@gmail.com)

Juan M. Carrillo-de-Gea
(University of Murcia, Spain,
https://orcid.org/0000-0002-3320-622X, jmcdg1@um.es)

Jaber El Bouhdidi
(Abdelmalek Essaadi University, Tetouan, Morocco,
https://orcid.org/0000-0003-3801-9316, jaber.elbouhdidi@uae.ac.ma)

Yacine El Younoussi
(Abdelmalek Essaadi University, Tetouan, Morocco,
https://orcid.org/0000-0003-4081-8248, yacine.elyounoussi@uae.ac.ma)

José L. Fernández-Alemán
(University of Murcia, Spain,
https://orcid.org/0000-0002-0176-450X, aleman@um.es)

Abstract: Due to the substantial development of information and communications technology, the use of E-learning in higher education has become essential to boost teaching methods and enhance students' learning skills and competencies. E-learning in Software Engineering turns out to be increasingly interesting for scholars. In fact, researchers have worked to enhance modern Software Engineering education techniques to meet the required educational objectives. The aim of this article is to analyse the scientific production on E-learning Software Engineering education by conducting a bibliometric analysis of 10,603 publications, dating from 1954 to 2020 and available in the Scopus database. The results reveal some scientific production information, such as the temporal evolution of the publications, the most prolific authors, institutions and countries, as well as the languages used. Besides, the paper evaluates additional bibliometric parameters, including the authors' production, journal productivity, and scientific cooperation, among other bibliometric parameters. The subject of the current study has not been treated by any previous bibliometric studies. Our research is deeper and more specific; it covers a long period of 66 years and a large number of publications, thanks to the chosen search string containing the different spellings of the used terms. In addition, the literature is analysed using several tools such as Microsoft Excel, VOSviewer, and Python. The research findings can be

used to identify the current state of E-learning Software Engineering Education, as well as to identify various research trends and the general direction of E-learning research.

Keywords: E-learning, Software, Education, Engineering, Bibliometric analysis, Scopus
Categories: K.3, K.3.1
DOI: 10.3897/jucs.87550

1 Introduction

In today's world, higher education systems are advancing due to the new tools of Information and Communication Technologies (ICT) [Gupta, 17][Mathivanan, 21]. They have had an enormous influence on the world economy, business management and new trends of globalization. Moreover, ICTs offer a potential to change the concept of studying worldwide, both for conventional and distance learning institutions [Palvia, 18]. E-learning systems keep on being a key role in our community since they help students in the educational process, making it possible for learners to be knowledgeable [Razzaque, 20].

People have started using computers and systems as a portion of learning processes since the sixties [Newbigcing, 59]. From then until now, this way of learning has increased. According to a study from the US Department Education, the attendance of on-line courses grew 65% [Aparicio, 14].

The integration of education and technology has driven the adoption of E-learning. Moreover, it has become a powerful medium for learning [Li, 20], and mainstream in the education sector, being adopted in higher education as well [Mathivanan, 21]. For researchers, E-leaning is a great research venue that has attracted attention [Recke, 21].

E-learning is based on an electronic environment. The different issues and divergent objectives, among which E-learning is used, have led to a disconnected understanding of how this type of system should be defined [Singh, 19]. The current status of the issues is probably best elucidated by a myriad of terms; the purpose of which is to refer to the progress of education delivered through computer technology. Some of these terms are E-learning, distance learning, virtual education, computer-assisted instruction, computer-based instruction [Elfaki, 19].

One of the advantages of E-learning is that it makes the learning process more enjoyable and free, as people can learn any subject by themselves, by simply accessing pre-recorded lessons through the use of technology. Since the users have the ability to access lessons at any time they prefer, the problem of missing a lesson no longer matters [Liu, 20]. The practice of E-learning skills, makes a remarkable impact on the process improvement, completeness, evolution, amendment and dynamic learning practice. The form of learning in this case is more interactive than the classical classroom approach, presenting the opportunity to manipulate learning methods to suit user preferences [Fatima, 19].

E-learning has been handled very successfully at the university, driving simple teaching and management strategies. Students can visit a variety of websites where they can learn concepts related to academic objectives. In addition, they can generate online meetings to share discussion with their classmates. Therefore, this web strategy makes the educational process more creative and flexible [Ayu, 18].

Recently, due to the COVID-19 pandemic, educational institutions around the world switched from classical teaching techniques to online sessions through

technology [Chang, 20]. Therefore, the COVID-19 crisis has shown the importance of E-learning in universities in order not to delay the teaching process [Soni, 20]. In fact, a recent study emphasizing education management during the COVID-19 crisis indicated that 90% of schools chose to use educational software techniques, 72% used online classes, and 40% provided links to follow up with online materials [Emami, 20].

Today, the training of software engineers is relevant, considering the importance of the software industry in which it is expected that there will be a deficit of software engineers. The barriers to training new software engineers do not end with scripting alone; on the contrary, students should be inspired to explore different ways of approaching Software Engineering. Thus, that they can understand which tools and techniques make sense to use in each case [Pinto, 17]. In this sense, software engineers must be able to apply a wide variety of methods in practice. Therefore, engineering educators face a challenging problem; mainly how to beneficially educate students about the entire software process. The use of technology with practical educational exercises can create an educational atmosphere with innovations that enhance study. However, the type of assignments offered in the classroom is restricted in the scope and duration, which increases the number of obstacles to find an equitable balance between theory and practice in lectures [Bosman, 20]. Technological advancement has produced differences in the learning styles of engineering students and the classical teaching of engineering teachers [Torres Kompen, 19]. Researchers have devoted efforts to improve and establish modern teaching methods in Software Engineering [Aloia, 19], as the Software Engineering education system must provide students with tailored materials to achieve the necessary educational goals [Lin, 19].

Bibliometrics is best known as a well-established form of research in Information Science, which allows the evaluation of research performance [Linnenluecke, 20]. To this end, it adopts quantitative analysis and statistics to analyse the bibliometric characteristics of a given domain, assessing the performance of writers/schools/nations discovering hot topics in an academic field, and revealing research trends [Yu, 20].

So far, many studies have analysed specific topics under a bibliometric approach. This paper is a bibliometric study on the current situation of E-learning systems use in Software Engineering education. The aim of this paper is to show an overview of the past, recent and upcoming trends in these research topics. The results of this paper can help to promote future research on the use of E-learning systems when teaching Software Engineering. In this bibliometric study, publications covering the period 1954-2020 were considered. As a result, this work presents various descriptive analyses, such as the temporal evolution of the publications, the most prolific authors and institutions, the geographical distribution of the publications, the languages used in the publications and the type of publications most relevant to the topic. In addition, the paper analyses the authors' production, journal productivity, scientific cooperation, authors' citation analysis and journal citation analysis, among other bibliometric parameters.

This article is organized as follows: Section 2 presents the related work, Section 3 describes the bibliometric techniques employed, the data source used in the analysis, and the process of constructing the search string, Section 4 addresses our results and discussion, and finally, Section 5 presents our conclusions and future work.

2 Related Work

Bibliometric methodologies are considered useful to support researchers in decision-making in the setting of research priorities and allows them to track the evolution of science. Bibliometric analysis measures the relevance of literature on a particular topic due to the statistical approach and quantitative analysis used. It is particularly helpful for creating a thorough overview of the top trends in a subject of study. Currently, E-learning is increasingly seen as the subject of many bibliometric overviews due to the growing interest given to the field [Tibaná-Herrera, 18]. In higher education, many researchers have created a literature overview on E-learning. Moreover, previous bibliometric studies have provided different perspectives of E-learning.

Djeki [22] conducted a bibliometric analysis of 12,272 publications between 2015 and 2020 from the Web of Science (WOS) database to examine the E-learning research domain . The results of his research showed that the countries with the highest output in E-learning are the USA, Spain, England, and China. The most prolific author in this field is Tarhini, A. In fact, the host universities that gather E-learning research from around the world are British, American, and Chinese universities. The most represented journals in the research field are Computers in Human Behaviour, Computer and Education, and International Journal of Emerging Technologies in Education. The most influential universities are the Universidad Nova de Lisboa, King Abdulaziz University and Islamic Azad University. In fact, according to the study, there is a limited collaboration between writers, universities, and countries working on E-learning, and COVID-19 has had a big impact on it.

Gao [22] proposed a bibliometric analysis of 1,985 journal articles gathered from the Scopus database. The documents were analysed using Biblioshiny and VOSviewer software. The results showed that publications on E-learning increased regularly with a 26.48% annual growth rate. “Computer and Learning” was considered the most impactful journal based on total citations. The “International Journal of Emerging Technologies in Learning” was the most productive. The quantity and total citations were published mostly by Europe and the United States. At the researcher level, Al-Samarraie, H. was considered the top author in both productivity and h-index, and Ebner, M. was the most influential author by total citations. The most often used keywords drawn from the cluster analysis were “distance learning”, “technology acceptance”, “teachers”, “professional training and development” and “quality assessment of E-learning”.

A total of 25,330 scientific output on E-learning has been analysed by [Sónia Rolland, 21] on the Scopus database from 2000 until 2019. The results of the bibliometric analysis revealed that rates of publication are still increasing. A report was generated on the publications in the field, including the journals, languages, authors, keywords, organizations, and countries. Computers And Education and International Journal of Emerging Technologies in Learning are the journals that published a higher number of publications related to the field. The bibliometric analysis reveals that Hwang, G. J., is the most prolific author. For more than 20 years, the United States has been considered the leader in terms of the number of articles. The Open University of the United Kingdom is the organization with the most references. The most commonly used language in publications is English.

A worldwide bibliometric study of the main improvements that have been published in E-learning was presented by [Das, 21] between 1910 and 2020. The

bibliometric data were restored from the Scopus database. The review contains the most cited, prolific writers and the fundamental schools and countries of the journal, based on bibliometric sign. In fact, the US, UK, China, and Australia were the world's leading nations in publishing articles on E-learning based on an analysis of 27,979 publications. In addition, Li, Y. was ranked the most prolific author with 82 articles. Moreover, one of the most widely cited authors in this field is Ahmad, R. He was selected with 2,472 citations. "E-learning", mobile learning" and "online learning" were the most commonly used keywords in articles of E-learning. The bibliometric analysis and the theoretical aspect have announced the idea of the importance of rebuilding and reproducing the journal by taking into consideration the publication citations, Scientometric performance, and subject covered. The study prepares a clear vision for the researchers to apply in order to achieve better results in the next decade.

Azurduy [21] presents a bibliometric analysis of the adoption of E-learning in higher education. This study selected 414 publications between 2006 and 2021. The results consider four thematic groups, a list of the 10 most cited articles, a morpheme cloud for ranges of study, and two theoretical aspects (citation and keywords). As a result, this paper represents the most influential authors, documents, key themes, current and future research fields. The findings show that Malaysia, the USA, the UK, and Indonesia provide the most academic publications on E-learning in higher education. The terms "e-learning", "online learning", "mobile learning", "blended learning", and "machine learning" are the most frequently used keywords. This paper can be considered as a reference for future research and as an illustration of how to perform a basic bibliometric analysis.

Fatima [19] examines the trends in the field of E-learning. A total of 9,826 publications covering the years 1989 to 2018 were gathered from the WOS database. The results showed that literature has grown significantly. Besides, half of the research output was produced by the US and the UK. The two journals that were most frequently used were Computers & Education and the Journal of Chemical Education.

An overall bibliometric analysis of 602 publications on E-learning in higher education that were published in the WOS database between 2020 and 2021 was provided by [Morales Muñoz, 22]. The dataset was analysed using VOSviewer, CiteSpace, and Knowledge Matrix Plus to extract networks and bibliometric indicators related to keywords, authors, organizations, and countries. The study ended up with different results within higher education. Actually, there are sub-fields of E-learning in higher education, such as distance learning, interactive learning, online learning, virtual learning, computer-based learning, digital learning, and blended learning. The prominent authors in this field are Moreno-Guerrero, A. J., and Lüftenegger, M. The University of Toronto, Canada, is the most cited institution in this field. Undoubtedly, the United States is one of the leading countries in terms of publications and citations. The idea of this research is about pedagogical techniques, particularly E-learning and collaborative learning. The sub-fields of artificial intelligence, machine learning, and deep learning constitute new research directions for E-learning approaches for further analysis.

3 Research Methodology

Bibliometric analysis is a popular tool to quantitatively analyse academic literature published in a specific area [Ma, 16]. Using a wide range of indicators and methods, it allows researchers to discover the bibliometric characteristics of the literature, as well as to show patterns in publication trends and to assess trends in future research development [Li, 20]. This study was carried out following the methodology of measuring academic research, which was explained by [Andrés, 09]. The bibliometric analysis allowed depicting an overall picture of the publications on E-learning Software Engineering Education. The aim of the bibliometric analysis is to assess the importance and impact of literature on Software Engineering Education and analyse various features of research output. The research design used in the study includes a variety of descriptive analyses, such as the temporal evolution of the publications, the most prolific authors and institutions, the geographical distribution of the publications, the languages used in the publications, and the types of articles most pertinent to the issue. The identification of the bibliographic database(s) is the first step, given its function as the main information source of scientific publications.

The data used for this bibliometric analysis were collected using one of the most important bibliographic databases created by publisher Elsevier in November 2004, which is the Scopus database. This academic resource is the largest bibliometric database and one of the most widely used sources for citing bibliometric data [García-Berná, 19]. The present bibliometric study considers a period of 66 years, from 1954 to 2020.

Query words and boolean operators were used to search for scientific documents in the field of E-learning Software Engineering Education. E-learning and software engineering were selected as the keywords under study. The Boolean operator AND and OR were used to combine keywords to procure a more accurate result. Therefore, a search string related to E-learning Software Engineering education was built with words that had to appear in the title, abstract, and keywords sections of the papers. Accordingly, the following search string was used:

TITLE-ABS-KEY ((elearning OR e-learning OR ((distance OR virtual OR electronic OR web-based OR online OR on-line OR remote OR computer-based OR computer-assisted OR computer-aided) AND (education OR learning OR teaching))) AND (("software engineering") OR ("software process") OR ("software requirements") OR ("software testing") OR ("project planning") OR ("project assessment") OR ("software risk") OR ("software configuration") OR ("software design") OR ("software construction") OR ("software implementation") OR ("software integration") OR ("software maintenance") OR ("software verification") OR ("software validation") OR ("software metrics")))

The results of the database search were exported for analysis in a bibliometric manner. The number of documents retrieved according to the defined query was 10,603 publications that were downloaded on February 9, 2021 in a CSV file. The collected data was processed in order to obtain the parameters of the bibliometric study.

The literature was analysed using several electronic tools to evaluate the data. Firstly, Microsoft Excel allowed checking visually that the data was properly downloaded. Moreover, the software library python data analysis library (Pandas,

https://pandas.pydata.org/) was employed to work with the data and calculate the parameters. In addition, this analysis is enriched with the visualization performed by VOSviewer software tools to present network relationships for a better presentation effect [García-Berná, 19].

The main objective of this research was to find the distribution of publications by year, most prolific institutions, geographical distribution of the publications, language used in publications, kind of publications and common keywords in the publications, most prolific authors, most preferred source, and scientific collaboration, among other aspects.

The bibliometric analysis provides a picture of the scientific research, which will help in descriptions, comparisons, and visualizations based on evidence of the results. The findings in this paper are used to describe briefly the impact of research and can be a guide for researchers who wish to know and contribute to the development and strengthening of the discipline.

4 Results and Discussion

4.1 Descriptive analysis

This section of the descriptive analysis explores four different parameters that are divided into four categories: (a) Temporal evolution with parameters such as growth of publications, annual growth rate (AGR) of the publications, compound annual growth rate (CAGR) of the publications. (b) Institution and countries with most prolific institutions and geographical distribution of the publications. (c) Language, in which the languages used in publications are shown. Finally, (d) type of document, which consists of the bibliometric variables such as kind of publication and common keywords in the publications.

4.1.1 Temporal evolution

Growth of publications

By observing Table 1, a significant increase in the number of publications in the field of E-learning Software Engineering Education is detected since 2009. The highest number of publications is 834, published in 2019. On the other hand, the lowest number of publications was found in 2010 with 321 manuscripts. There is an increasing trend concerning the growth literature in the topic.

This progress could be attributed to the successful integration of Internet resources with education to facilitate teaching-learning processes. As a result, an increasing number of organizations, schools, and enterprises have begun to provide E-learning courses in recent years [Kasani, 19].

Engineers, academics, and researchers are continually searching for new pedagogical approaches and procedures to improve Software Engineering education. They are mainly concerned with improving the educational process based on E-learning [Asgariid, 21]. Additionally, in the beginning of 2020, the COVID-19 outbreak has forced the universities around the world to adopt online learning due to its beneficial properties [Amin Almaiah, 20]. E-learning encourages lifelong learning and self-

directed learning which is not restricted by the gender or location of learners. Moreover, students can access the instructional materials from anywhere at any time, 24 hours a day, 7 days a week [Talaat Abumandour, 18]. All the mentioned factors may explain the development of the number of publications in the field of our research.

Year	Publications	%
2020	768	7.24
2019	834	7.87
2018	657	6.20
2017	606	5.72
2016	538	5.07
2015	450	4.24
2014	470	4.43
2013	405	3.82
2012	397	3.74
2011	473	4.46
2010	321	3.03
<2009	4684	44.18

Table 1: Growth of publications

Annual Growth Rate of publications (AGR)

The AGR represents the total number of publications generated in comparison to the previous year (2). The number of publications in one year and the number of publications from the previous year are used to determine this factor. Table 2 depicts the AGR of literature publishing throughout the period 2010-2020. The highest AGR was 47.35 % in 2011, followed by 31.55 % in 2010 and -16.06 % in 2012. The following formula (1) was used to compute this parameter.

$$\mathrm{AGR}(\%) = 100 \cdot \frac{\mathrm{NumberOfPublications}_{\mathrm{year}} - \mathrm{NumberOfPublications}_{\mathrm{year-1}}}{\mathrm{NumberOfPublications}_{\mathrm{year-1}}} \quad (1)$$

Year	Percentage
2020	-7.91 %
2019	26.94 %
2018	8.42 %
2017	12.64 %
2016	19.56 %
2015	-4.26 %
2014	16.05 %
2013	2.02 %
2012	-16.07 %

2011	47.35 %
2010	31.56 %

Table 2: Annual growth of the publications (AGR)

Compound annual growth rate of the publications (CAGR)

The CAGR is the rate of return required for an investment to increase from its initial balance to its final balance, assuming profits are reinvested at the end of each year of the investment's life cycle. The compound annual growth rate was determined in this study by taking the nth root of the total percentage growth rate, where n is the number of years in the period under consideration. This can be expressed in the following way (2):

$$\mathrm{CAGR}(\%) = 100 \cdot \left[\left(\frac{\mathrm{CumulativeNumberOfPublications}_{\mathrm{year}}}{\mathrm{NumberOfPublications}_{\mathrm{year}}} \right)^{\frac{1}{\mathrm{year}-\mathrm{referenceYear}}} - 1 \right] \quad (2)$$

The compound annual growth rate of publication on the topic is seen in Table 3 and Figure 1. The highest CAGR of 73.20 % was registered in 2012, followed by a CAGR of 67.86 % in 2011.

Year	Cumulative	Percentage
2020	5919	22.65 %
2019	5151	22.42 %
2018	4317	26.53 %
2017	3660	29.29 %
2016	3054	33.56 %
2015	2516	41.09 %
2014	2066	44.79 %
2013	1596	57.95 %
2012	1191	73.20 %
2011	794	67.86 %
2010	321	-

Table 3: Compound annual growth rate (CAGR)

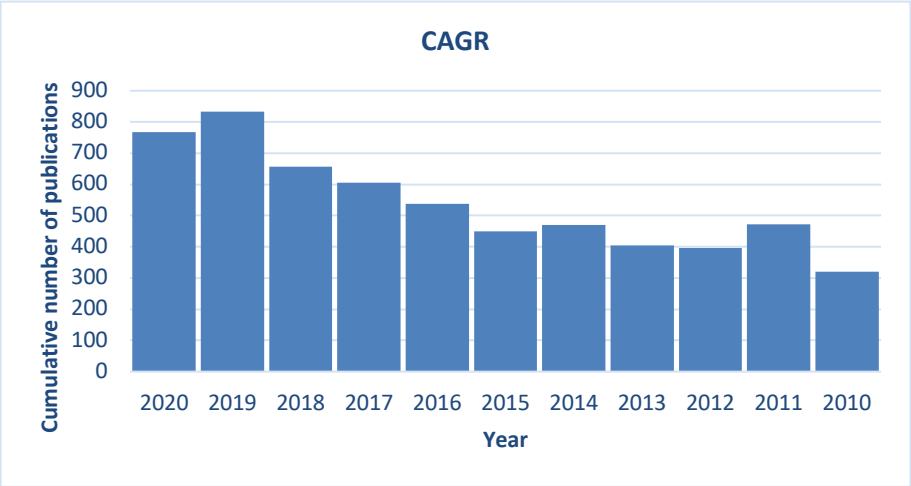


Figure 1: Compound annual growth rate (CAGR)

Relative Growth Rate (RGR) and Doubling Time (DT)

The RGR and DT of publication from the selected period of research are depicted in Table 4 and Figure 2. The highest value of RGR was 0.90 in 2011, followed by the lowest value of 0.40 in 2012. The RGR and DT models were used to calculate the growth rate of all publications. The DT is directly related to RGR. Moreover, it is used to assess the rise in the number of publications on time. The mathematical representation of the mean relative growth rate of papers over a specific period of time is specified with the following formula (3):

$$\mathrm{RGR} = \frac{\mathrm{Ln}(\mathrm{W}_2) - \mathrm{Ln}(\mathrm{W}_1)}{year_2 - year_1} \quad (3)$$

It is determined that the RGR and DT have a direct equivalence based on the calculations. If the number of contributions made by a subject doubles during the course of the research, the difference between the logarithms of the numbers at the beginning and end of the time must be the logarithms of the number 2. This difference has a value of 0.693 [Beaie, 09]. The following is the formula of comparable DT for contributions and pages measurement when using a Neper logarithm (4):

$$\mathrm{DT} = \frac{\mathrm{Ln}(2)}{\mathrm{RGR}} \quad (4)$$

In 2020, the highest DT of 4.98 was obtained, followed by 4.19 in 2018. Table 4 shows the overall statistics for RGR and DT.

Year	Cumulative	Ln(W1)	Ln(W2)	RGR	DT (Years)
2020	5919	8.54	8.68	0.13	4.98

2019	5151	8.37	8.54	0.17	3.92
2018	4317	8.20	8.37	0.16	4.19
2017	3660	8.02	8.20	0.18	3.82
2016	3054	7.83	8.02	0.19	3.57
2015	2516	7.63	7.83	0.19	3.51
2014	2066	7.37	7.63	0.25	2.68
2013	1596	7.08	7.37	0.29	2.36
2012	1191	6.67	7.08	0.40	1.70
2011	794	5.77	6.67	0.90	0.76
2010	321	8.68	5.77	0.29	2.37

Table 4: Relative growth rate (RGR) and Doubling time (DT)

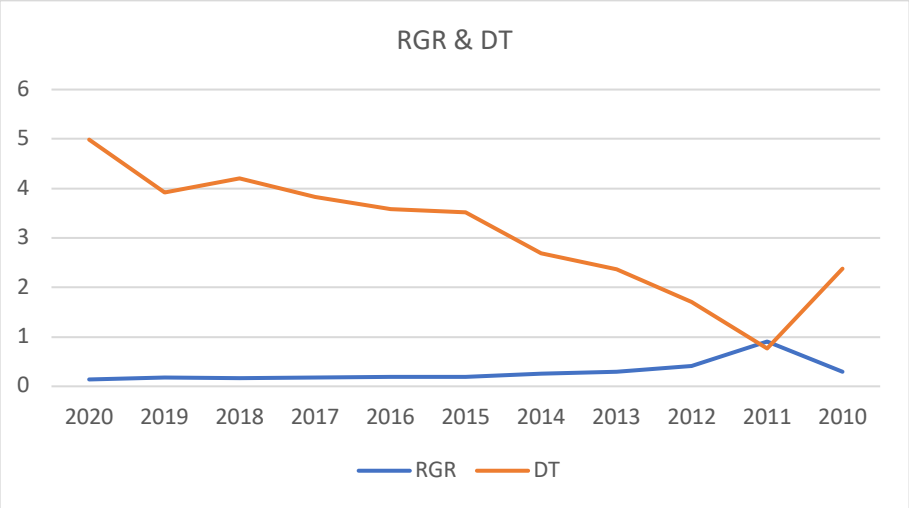


Figure 2: Relative growth rate (RGR) and Doubling time (DT)

Trend analysis for the number of publications

The least squares method is used to make a trend analysis in order to estimate the number of publications that may emerge in the future. The following formula yields a straight line: f(X) = b*X + a, which is determined with data from the previous decade, from 2010 to 2020. This formula was obtained by solving the system of equations (5) and (6), where N is the number of years under consideration, Y denotes the estimated number of publications each year, and X denotes a conveniently selected input of the equation. To make it easy to determine the coefficients, X values are chosen so that when all of the values are added up, they equal zero. As a result, the coefficients' values are the ones shown in (7) and (8), which are obtained from the values in Table 5.

$$\sum Y = N \cdot a + b \cdot \sum X \qquad (5)$$

$$a = \frac{\sum Y}{\mathrm{N}} = \frac{5151}{10} = 515.1 \qquad (7)$$

$$\sum X \cdot Y = a \cdot \sum X + b \cdot \sum X^2 \quad (6) \qquad b = \frac{\sum X \cdot Y}{\sum X^2} = \frac{7329}{330} = 22.2 \quad (8)$$

Year	Actual number of publications (Y)	Input (X)	X·Y	X2	Number of publications trend f(X)	Difference
2024		19	0	361	937.07	44.41
2023		17	0	289	892.65	44.41
2022		15	0	225	848.23	44.41
2021		13	0	169	803.81	44.41
2020	768	11	8448	121	759.4	-66
2019	834	9	7506	81	714.98	177
2018	657	7	4599	49	670.56	51
2017	606	5	3030	25	626.14	68
2016	538	3	1614	9	581.72	88
2015	450	1	450	1	537.30	-20
2014	470	-1	-470	1	492.89	65
2013	405	-3	-1215	9	448.47	8
2012	397	-5	-1985	25	404.05	-76
2011	473	-7	-3311	49	359.63	152
2010	321	-9	-2889	81	315.21	-
$\sum_{n=2010}^{2019}$	5151	0	7329	330		

Table 5: Computation of straight-line trend using the least squares method

Table 5's "Difference" column shows the separation of the number of publications between two consecutive years. This difference has been estimated from 2010 to 2020, using real values collected from Scopus database (column Y). Nevertheless, starting from 2021, it is only possible to calculate the difference with estimated values (column f(X)). In comparison to the previous year, the straight-line calculation predicts that around 44 extra publications will be created each year. Figure *3* 3 illustrates the relationship between the number of publications published each year and the actual number of publications published each year.

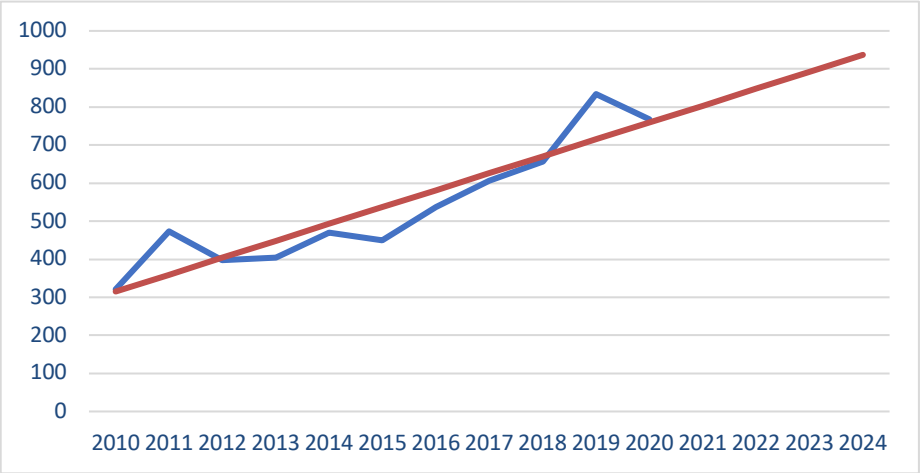


Figure 3: Tendency analysis graph of the number of publications per year

4.1.2 Institutions and countries

Most prolific institutions

The most productive institutions are shown in Table 6. According to the results of the study, the most productive and prolific institution related to E-learning Software Engineering education is the University of California, which has 78 papers, accounting for 0.81% of all publications in this discipline. This institution is followed by the Carnegie Mellon University, which has 69 publications (0.71%), Purdue University, which has 58 publications (0.60 %), Arizona State University, and Georgia Institute of Technology, which both have 56 publications. This can be explained by the importance that the university gives to E-learning since the beginnings of online education. It was one of the first universities to adopt online learning [Kentnor, 15]. The California Virtual University, a consortium of almost 100 universities and colleges in California with nearly 1,600 online courses, opened in November 1998 [Karen, 98]. This may encourage the university's research in this area to improve online learning. Currently the University of California has provided more than 100 programs to 4 million students from 150 countries and regions since joining the EDX Platform partnership in 2012. University of California, Berkeley MOOCs are also available with online courses in a variety of disciplines [Pappano, 12]. The achievements of the University of California in the field of E-learning certainly continues to develop thanks to the research and the interest given to the field.

Geographical distribution of the publications

Table 7 and Figure 4 shows the most productive countries in terms of research publications in the field of E-learning Software Engineering Education. It can be observed that the USA is the country with the most publications, with 2,976 manuscripts, representing 28.06% of total publications, far more than any other country. Following the United States, China and the United Kingdom are ranked second

and third, respectively, with 993 and 643 publications. Many other countries, such as Germany, Spain, Canada, Australia, Italy, Brazil, India, Japan, and France, also generated notable results.

Pos.	Institution	Number of publications	Percentage
1	University of California	78	0.81 %
2	Carnegie Mellon University	69	0.71 %
3	Purdue University	58	0.60 %
4	Arizona State University	56	0.58 %
5	Georgia Institute of Technology	56	0.58 %
6	Washington University	50	0.52 %
7	Complutense University of Madrid	45	0.47 %
8	Pennsylvania State University	45	0.47 %
9	North Carolina State University	42	0.43 %
10	University of Pittsburgh	40	0.41 %

Table 6: Most prolific institutions

This can be the result of the great technological advances, and the great interest in distance education in USA and China. In the USA over 77% of US businesses in 2017, used online learning, but by 2020, 98% intended to include it in their programs [Chernev, 22]. Around 65% of faculty members in the United States support online learning courses and educational materials [Johnson, 20]. Moreover, in China E-learning continues to grow. Major research areas such as student's and teacher's perspectives on developments in E-learning, teacher's pedagogical capacity, and the production of more convenient and useful E-learning resources are likely to be topics of continuing research interest [Wang, 18]. The spread of the use of E-learning systems in education in China may be due to the development of internet usage. According to the statistical report, the number of internet users in China has reached 0.731 billion in December 2015, while the number of networked people using mobile phones reached 0.695 billion [Cnnic, 17]. China's Internet penetration rate has reached 53.2%, with more than half of the population connected [Wang, 18]. E-learning in Europe is not widely accepted as it is in USA and China, but according to figures issued by the EU Statistical Office, the number of European Union nationals and residents taking online courses climbed by 4 percentage points in 2021 compared to previous years [Eurostat, 22]. Thus, it can be guessed that research in Europe in the field of E-learning Software Engineering education will flourish in the years to come, given the remarkable interest shown in this field [Jokiaho, 18].

Pos.	Country	Number of publications	Percentage
1	United States	2,976	28.06 %
2	China	993	9.36 %

3	United Kingdom	643	6.06 %
4	Germany	580	5.47 %
5	Spain	509	4.80 %
6	Canada	370	3.48 %
7	Australia	321	3.02 %
8	Italy	270	2.54 %
9	Brazil	268	2.52 %
10	India	249	2.34 %
11	Japan	242	2.28 %
12	France	189	1.78 %

Table 7: Geographical distribution of publications

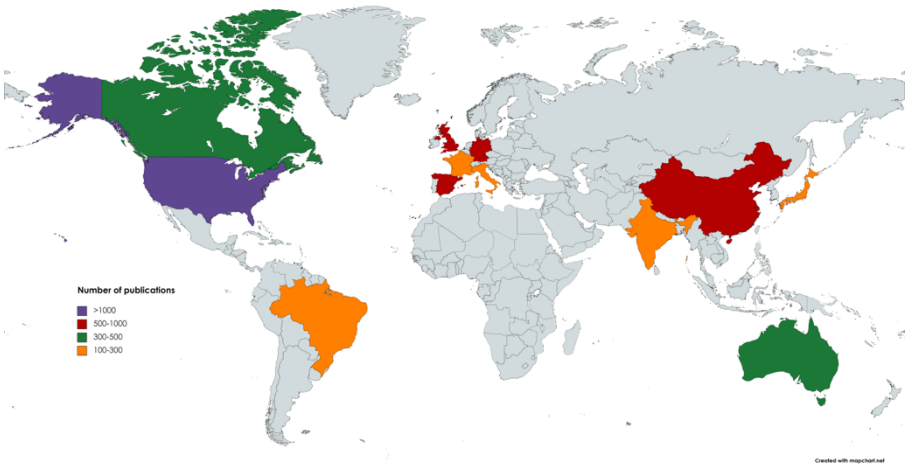


Figure 4: Most productive countries

4.1.3 Languages

Language used in publications

Table 8 indicates the languages used in the publications related to E-learning Software Engineering education. English is the most common, and then a number of publications between 13 and 54 were found in other languages such as Chinese, Spanish, German, Portuguese Turkish and Russian. At a lower level, with less than 7 publications French, Japanese and Ukrainian were found.

This can be explained by the fact that English is widely used in the field of science and technology. According to the International Federation on Documentation nearly 85% of all the scientific and technological information in the world is written and/or abstracted in English [Ammon, 01].

It has also become the de facto universal language, with significant implications for scientific communication. Consequently, scientists from all over the world may access available scientific information and communicate with scientists from different

locations, regardless of where they are located. It is common knowledge that English is used extensively in the sphere of education across the world [Ferguson, 11], because the majority of higher education literature is published in English. In this vein, learning English has become a need. English is extensively used by students, professors, and researchers all over the globe, since it is necessary in many sectors of education and is considered the language of information, which is kept in the form of books and journals in both written and electronic form [Houaria, 21].

Pos.	Language	Number of publications
1	English	10,381
2	Chinese	54
3	Spanish	51
4	German	49
5	Portuguese	29
6	Turkish	14
7	Russian	13
8	French	7
9	Japanese	6
10	Ukrainian	6

Table 8: Distribution as regards languages used

4.1.4 Type of document

Kind of publications

Table 9 shows the distribution of publications based on the channels employed. More than half of the publications on E-learning Software Engineering education are in the form of articles, according to the results. The total number of conference paper are 6,870, which means 64.79% over the total amount of publications in the field, followed by articles (26.72%), Conference Reviews (5.08%) and Reviews (1.51%).

Pos.	Language	Number of publications
1	Conference Paper	6870
2	Article	2833
3	Conference Review	539
4	Review	161
5	Book Chapter	121
6	Book	36
7	Editorial	15
8	Short Survey	11

9	Letter	6
10	Note	5
11	Article in Press	3

Table 9: Forms of publication

Common keywords in the publications

The most prevalent keywords found in the publications are listed in Table 10. These terms are "E-learning" and "Software Engineering" followed by "Education", "Machine Learning", "Simulation", "Software Engineering Education", "Virtual Reality", "Software Development", "Software Tools" and "Learning". All these keywords correspond to concepts, which are closely related to the field under study.

Among the top of the keywords list "Machine Learning" is the most appealing. This can be explained by the integration of machine learning algorithms in several applications. Nowadays, researchers can employ machine learning measures to tackle issues in the E-learning sector [Farhat, 20]. Learning management systems (LMS) may grow smarter and more responsive to users as their demands are better met. This may be accomplished by employing machine learning approaches such as quantitative data analysis to examine student performance and supervised classification to do tasks, like sentiment analysis and student style categorization [Moubayed, 18].

Pos.	Keyword	Number of publications
1	E-learning	451
2	Software engineering	446
3	Education	325
4	Machine learning	265
5	Simulation	166
6	Software engineering education	133
7	Virtual reality	122
8	Software development	122
9	Software tools	109
10	Learning	109

Table 10: Keywords most frequently referred to

Figure 5 shows a label map of the main keywords that appeared in Scopus database. The label is more important if the text and the surrounding circle are larger. The most often used keywords in this illustration are E-learning, Education, and Machine Learning.

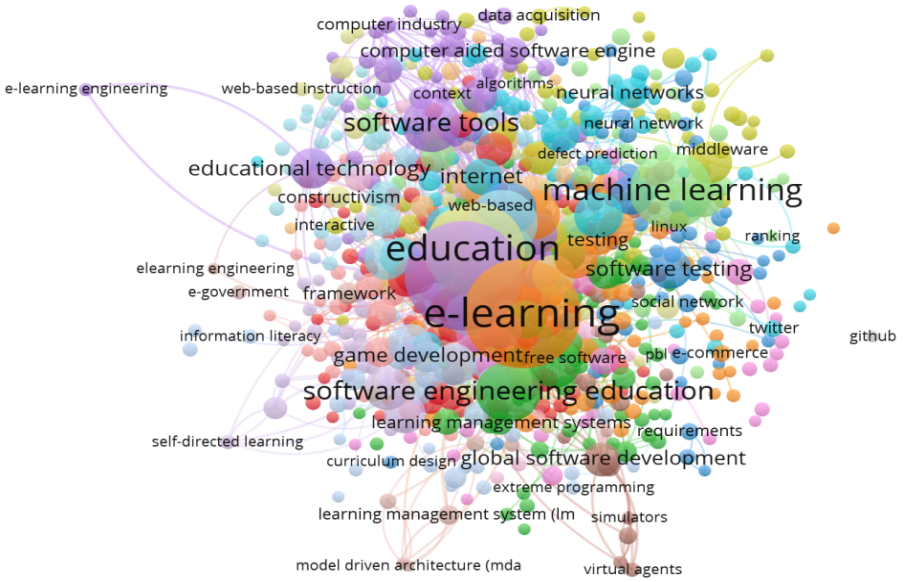


Figure 5: Most frequently used keywords

4.2 Author production

In this section, the parameter "number of authors" is presented, in which parameters such as author participant's productivity, authorship trend analysis, and most prolific authors are studied.

4.2.1 Number of authors

Author participants' productivity

An estimation of the mean author participant's productivity was calculated using the equations (9) and (10). Table 11 displays the results of an estimate of the average number of authors per paper (AAPP) and the average productivity per author (PPA).

$$\text{AAPP} = \overline{\text{AuthorsPerPaper}} = \frac{\text{NumberOfAuthors}}{\text{NumberOfPapers}} \quad (9)$$

$$\text{PPA} = \overline{\text{ProductivityPerAuthors}} = \frac{1}{\overline{\text{AuthorsPerPaper}}} \quad (10)$$

Years	Papers	Authors	AAPP	PPA
2020	768	2862	3,73	0,27
2019	834	3008	3,60	0,28
2018	657	2238	3,41	0,29
2017	606	2039	3,36	0,30

2016	538	1776	3,30	0,30
2015	450	1502	3,34	0,30
2014	470	1472	3,13	0,32
2013	405	1233	3,04	0,33
2012	397	1153	2,90	0,34
2011	473	1370	2,90	0,35

Table 11: Author productivity

From the initial years, when publications on E-learning Software Engineering education appear, until 2020, where the value of AAPP is around 4, the findings (Table 11) indicate that it is usual to find papers produced by 2 or 3 authors. Moreover, it seems to be a tendency toward more cooperation. On the other hand, across the research period, the mean value of PPA decreased from 0.38 to 0.26.

Authorship trend analysis

Table 12 shows the total number of publications each year divided by the number of writers that collaborated on each one. Instead of using the entire number of publications retrieved from the Scopus database, the percentages were computed using the total amount of publications acquired by adding up the publications that appeared from 2011 to 2020 (5,598 in Table 12). During the last decade and according to the results in Table 12, the literature regarding E-learning Software Engineering education was mostly written by multiple authors. From 2011, multiple author's publications have experienced an upward and kept on growing till the year this research was conducted.

Year	Single author		Multiple authors		Number of publications
	Total output	**%**	**Total output**	**%**	**Total output**
2020	58	1.03	675	12.05	768
2019	79	1.41	724	12.93	834
2018	60	1.07	553	9.87	657
2017	67	1.19	517	9.23	606
2016	57	1.01	459	8.19	538
2015	54	0.96	377	6.73	450
2014	68	1.21	373	6.66	470
2013	42	0.75	338	6.03	405
2012	50	0.89	319	5.69	397
2011	81	1.44	376	6.71	473
TOTAL	616	11.00	4,711	84.15	5,598

Table 12: Authorship trend analysis

Most prolific authors

Table 13 presents the results of determining the most prolific authors in the field of E-learning software engineering education. The most prolific author is Piattini, who has 18 publications, followed by Vizcaíno, who has 16 publications, and Barbosa and

Monasor with 14 publications. The percentages were calculated based on the total number of papers found through in the Scopus search.

The top author in the research topic based on total papers during the analysed period and total citations is Mario Piattini. He is a computer scientist in the area of systems and software engineering. He founded the Alarcos Research Group of the University of Castilla-La Mancha (Spain), where he also served as deputy director of the Department of Computer Science during the early 2000s. He also founded the UCLM-INDRA Joint Research and Development Center, where he was director until 2010. Meanwhile, he founded the Institute of Technologies and Information Systems (ITSI) of the University of Castilla-La Mancha, where he was director until 2015. He was identified among the fifteen best researchers in the world in the field of systems and software engineering in an independent study (2004-2008). This may explain why he is the first on the list of most prolific authors in the field of E-learning Software Engineering education.

Pos.	Author	Number of publications	Percentage
1	Piattini, M.	18	0,17
2	Vizcaíno, A.	16	0,15
3	Anonymous	14	0,13
4	Barbosa, E.F.	14	0,13
5	Monasor, M.J.	14	0,13
6	Bagert, D.J.	12	0,11
7	Dormido, S.	12	0,11
8	García-Peñalvo, F.J.	12	0,11
9	Virvou, M.	12	0,11
10	Clarke, P.I.	11	0,10
11	Krusche, S.	11	0,10

Table 13: Most prolific authors

4.3 Journal productivity

In this section, there is a more detailed analysis of the sources, such as journals and conferences. The journals and conferences are ranked based on the frequency of the documents published.

4.3.1 Journal and conference ranks

Most preferred sources

Table 14 indicates the percentage (column 4) that represents each source as regard the 10,603 publications retrieved from the Scopus database. The most preferred sources in which to publish can be divided into journals and conferences. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) (4.35%) are the most common sources in which it is possible

to find publications related to this area. Then, ASEE annual conference and exposition, conference proceedings (2.67%), and ACM international conference proceeding series (2.33%) were the conferences in which appeared the largest amount of literature related to this topic. The rest sources of publications in the top ten ranking, with a percentage of publications close to 1% are Proceedings-frontiers in education conference; ASEE annual conference proceedings; Proceedings-frontiers in education conference, FIE; CEUR workshop proceedings; Proceedings - international conference on computer science and software engineering, CSSE 2008; Communications in computer and information science and Proceedings-international conference on software engineering.

Pos.	Source	Number of publications	Percentage
1	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	462	4.35 %
2	ASEE Annual Conference and Exposition, Conference Proceedings	284	2.67 %
3	ACM International Conference Proceeding Series	248	2.33 %
4	Proceedings - Frontiers in Education Conference	195	1.83 %
5	ASEE Annual Conference Proceedings	177	1.66 %
6	Proceedings - Frontiers in Education Conference, FIE	171	1.61 %
7	CEUR Workshop Proceedings	145	1.36 %
8	Proceedings - International Conference on Computer Science and Software Engineering, CSSE 2008	132	1.24 %
9	Communications in Computer and Information Science	126	1.18 %
10	Proceedings - International Conference on Software Engineering	118	1.11 %

Table 14: Most preferred sources

4.4 Scientific collaboration

In this section, the collaboration between authors or institution has been analysed by means of indicators, such as the degree of collaboration, collaboration index (CI) and co-authorship index (CAI).

4.4.1 Collaboration Index

Degree of collaboration

Table 15 and Figure 6 denote that around 24% of the publications had 2 authors, and 21% had 3 authors. These figures are closed to the ones corresponding to 1 and 4 authors, which are 16% and 14% of the literature produced, respectively.

The degree of collaboration between the authors and the cooperation tendency were determined, using the result in Table 15. Moreover, equation (11) was used to calculate this factor.

Pos.	No. of authors	No. of papers	Percentage
0	Anonymous	553	5.21 %
1	Single author	1751	16.51 %
2	Two authors	2636	24.86 %
3	Three authors	2322	21.89 %
4	Four authors	1583	14.92 %
5	Five authors	815	7.68 %
6	Six authors	434	4.09 %
7	Seven authors	193	1.82 %
8	Eight authors	123	1.16 %
9	Nine authors	60	0.56 %
10	Ten authors	45	0.42 %

Table 15: Authorship pattern of publications

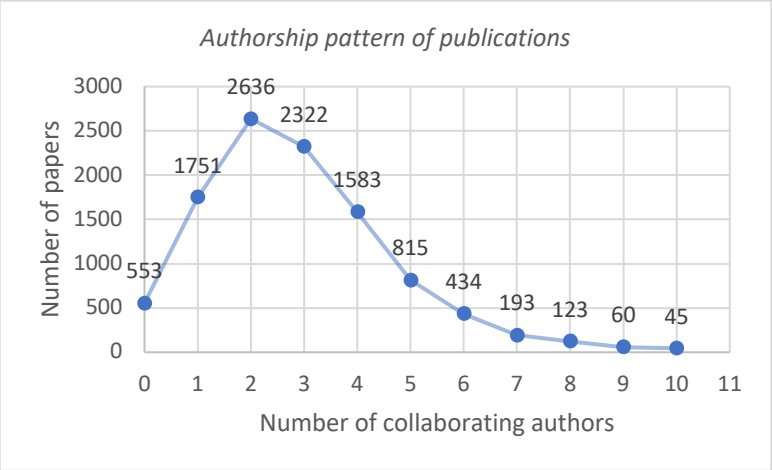


Figure 6: Authorship pattern of publications

$$C = \frac{\text{NumberOfPublications}_{\text{multipleAuthors}}}{\text{NumberOfPublications}_{\text{multipleAuthors}} + \text{NumberOfPublications}_{\text{singleAuthor}}} \quad (11)$$

The degree of collaboration was C = 0.8258. Thus, approximately 82% of the publications usually have more than one author. The value of the index decreased to 0.7915 after excluding publications for which no author was identified, indicating that authors in the subject of E-learning Software Engineering education tend to collaborate more. On both cases, the value is closed to 1, which means that the cooperation between authors is extended.

CI

The number of writers who contributed to the production of the papers is studied to determine a trend over time. This information is represented by the cooperation index (CI), which is calculated as given in the equation (12). The results are listed in Table 16.

$$\mathrm{CI} = \frac{\mathrm{NumberOfSignatoriesInMultiauthoredPublications}}{\mathrm{NumberOfMultiauthoredPublications}} \qquad (12)$$

As Table 16 shows, the value for CI remains around 3 over the years and reached the value of 4 in 2018. Consequently, a high collaboration pattern can be inferred. This pattern can also be inferred from the growth in both multi-authored publications and total authors in multi-authored publications parameters.

Year	Multi-authored publications	Total signatories in multi-authored publications	CI
2020	675	2862	4.24
2019	724	3008	4.15
2018	553	2238	4.04
2017	517	2039	3.94
2016	459	1776	3.86
2015	377	1502	3.98
2014	373	1472	3.94
2013	338	1233	3.64
2012	319	1153	3.61
2011	376	1370	3.64
2010	258	873	3.38

Table 16: Collaboration index

4.4.2 National and international collaborations

A collaboration profile can be obtained by calculating the number of publications that were published by multiple authors from different institutions and countries. The nature of these collaborations can be (1) international when authors are from different countries, (2) national when authors are from different institutions but belong to the same country, and (3) no collaboration in the remaining papers. Figure 7 represent these collaborations, and shows an increase in both national and international collaborations.

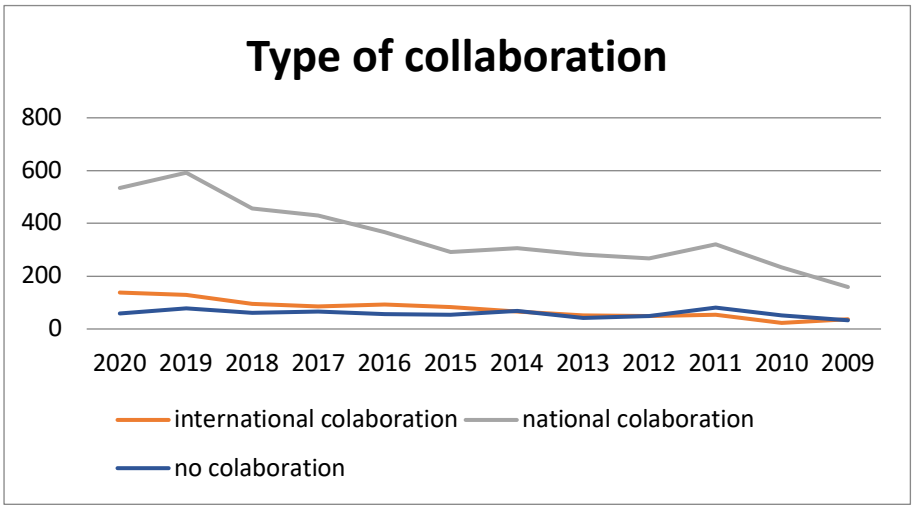


Figure 7: Collaboration pattern in international and international publications

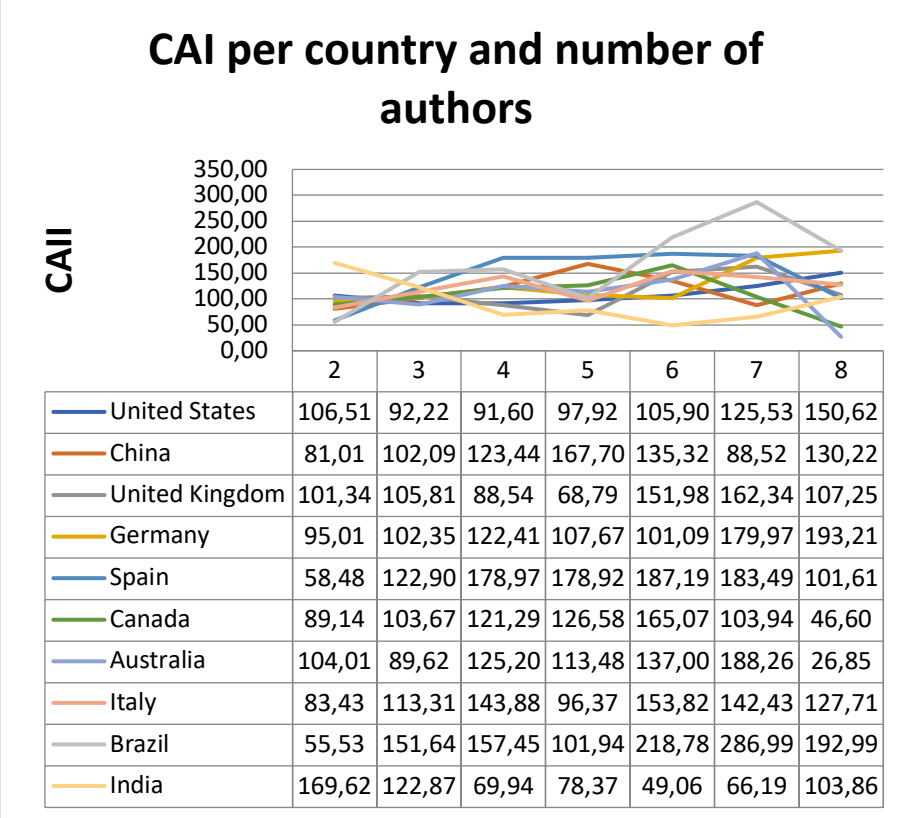
4.4.3 CAI

The co-authorship index (CAI) is a tool for determining if a country's number of publications matches the average within a co-authorship pattern [Andrés, 09]. The following equation (13) is used to determine this index:

$$CAI = ((N_{ca}/N_{ct})/(N_{ta}/N_{tt})) \times 100 \quad (13)$$

Where Nca is the total number of papers co-authored by authors in the cth country, Nct is the total number of papers in the cth country, Nta is the total number of papers co-authored by authors in the total number of countries, and Ntt is the total number of papers in the total number of countries.

The CAI values for co-authored papers (from 2 to 8 authors) in the 10 countries with the most publications are shown in Figure 8. Brazil has the highest CAI value for articles written with 6 and 7 authors, followed by Germany with 8 authors.



	2	3	4	5	6	7	8
United States	106,51	92,22	91,60	97,92	105,90	125,53	150,62
China	81,01	102,09	123,44	167,70	135,32	88,52	130,22
United Kingdom	101,34	105,81	88,54	68,79	151,98	162,34	107,25
Germany	95,01	102,35	122,41	107,67	101,09	179,97	193,21
Spain	58,48	122,90	178,97	178,92	187,19	183,49	101,61
Canada	89,14	103,67	121,29	126,58	165,07	103,94	46,60
Australia	104,01	89,62	125,20	113,48	137,00	188,26	26,85
Italy	83,43	113,31	143,88	96,37	153,82	142,43	127,71
Brazil	55,53	151,64	157,45	101,94	218,78	286,99	192,99
India	169,62	122,87	69,94	78,37	49,06	66,19	103,86

Figure 8: The co-authorship index plot

4.5 Author citation analysis

An author citation analysis will be conducted in this section in order to detect relationships between authors and sources, such as journals. The most commonly referenced articles and their authors will be found in this section.

4.5.1 Most cited publications

Table 17 shows the publications sorted by the number of citations in Scopus. The most frequently cited manuscripts are “Empirical Examination of the Adoption of WebCT Using TAM”, with 536 citations, and “Interaction, Internet Self-Efficacy, and Self-Regulated Learning as Predictors of Student Satisfaction in Online Education Courses”, with 400 citations.

Pos.	Title	Authors	Source	Citations
1	Empirical Examination of the Adoption of WebCT Using TAM	Ngai E.W.T.; Poon J.K.L.; Chan Y.H.C.	Computers and Education	536

2	Interaction, Internet Self-Efficacy, and Self-Regulated Learning as Predictors of Student Satisfaction in Online Education Courses	Kuo Y.-C.; Walker A.E.; Schroder K.E.E.; Belland B.R.	Internet and Higher Education	400
3	Asynchronous Learning Networks as a Virtual Classroom	Hiltz S.R.; Wellman B.	Communications of the ACM	382
4	Distance Learning Applied to Control Engineering Laboratories	Aktan B.; Bohus C.A.; Crowl L.A.; Shor M.H.	IEEE Transactions on Education	318
5	Socialization in an Open Source Software Community: a Socio-Technical Analysis	Ducheneaut N.	Computer Supported Cooperative Work: CSCW: an International Journal	305
6	An Instructional Model for Web-Based E-Learning Education with a Blended Learning Process Approach	Alonso F.; López G.; Manrique D.; Viñes J.M.	British Journal of Educational Technology	302
7	Using Structural Context to Recommend Source Code Examples	Holmes R.; Murphy G.C.	Proceedings - International Conference on Software Engineering	276
8	Experience with a Learning Personal Assistant	Mitchell T.M.; Caruana R.; Freitag D.; McDermott J.; Zabowski D.	Communications of the ACM	239
9	Issues in Using Students in Empirical Studies in Software Engineering Education	Carver J.; Jaccheri L.; Morasca S.; Shull F.	Proceedings - International Software Metrics Symposium	226
10	Clustering and Sequential Pattern Mining of Online Collaborative Learning Data	Perera D.; Kay J.; Koprinska I.; Yacef K.; Zaane O.R.	IEEE Transactions on Knowledge and Data Engineering	213

Table 17: Most cited publications

4.6 Journal citation analysis

In this step, the impact factor of the top 10 sources will be determined based on the number of papers obtained. In addition, a journal co-citation study will be conducted to detect links between journals and conferences.

4.6.1 Journal Impact factor

Table 18 depicts the most important sources in E-learning Software Engineering education, by displaying the values for the 2019 Scimago Journal Rank (SJR). By observing the results, it is noticeable that the most dominant type in the 10 most relevant sources is conferences, and there are only two journals that do not belong to the Journal Citation Report (JCR).

The results founds that Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) (4.35%) is the most common sources in which it is possible to find publications related to this area. The conferences with the most literature related to the used search string were the ASEE annual conference and exhibition, conference proceedings (2.67%), and ACM international conference proceeding series (2.33%).

4.6.2 Co-citation analysis

Figure 9 shows a network by using co-citation analysis in journals, thus identifying the most commonly co-cited groups of journals. Based on the mapping and clustering approach performed, a set of clusters were identified in the scientific literature landscape including Computers & Education, Communications of the acm, IEEE software and others.

Quartile SJR | JCR	Source	2019 SJR
Q2 | -	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	0.43
---	ASEE Annual Conference and Exposition, Conference Proceedings	0.24
---	ACM International Conference Proceeding Series	0.20
---	Proceedings - Frontiers in Education Conference	0.32
---	ASEE Annual Conference Proceedings	0.24
---	Proceedings - Frontiers in Education Conference, FIE	0.32
---	CEUR Workshop Proceedings	0.18
---	Proceedings - International Conference on Computer Science and Software Engineering, CSSE 2008	0
Q3 | -	Communications in Computer and Information Science	0.19
---	Proceedings - International Conference on Software Engineering	0.52

Table 18: 2019 Impact Factors

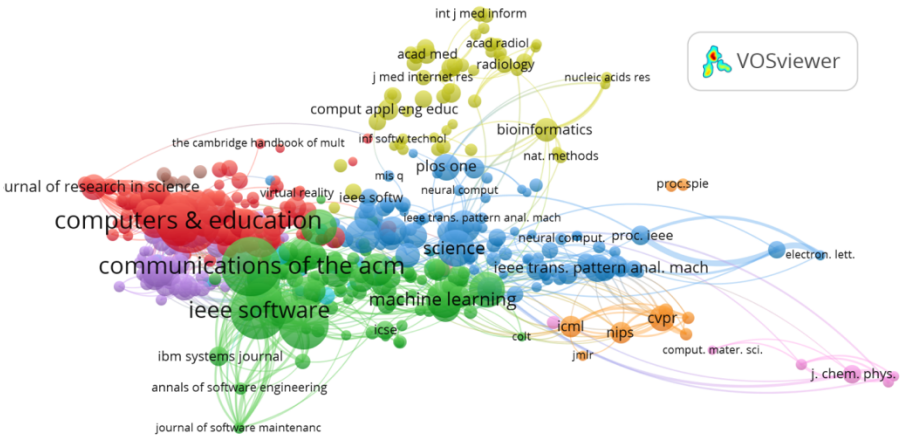


Figure 9: Network map revealing co-citation of journals

4.7 Comparative study

Table 19 presents a comparison of five bibliometric studies already cited in the related work section along with our research. We excluded [Sónia Rolland, 21] and [Fatima, 19], because the tools used to obtain the results of the bibliometric analysis were not mentioned in the papers. The majority of previous studies adopted the subject of E-learning in general, or E-learning in higher education. There is no bibliometric analysis of scientific literature that focuses on E-learning software engineering education.

Gao [22] and Das [21] collected their data from the same database used in our study, which is "Scopus". Azurduy [21] used The lent database and the other two authors [Djeki, 22] and [Morales Muñoz, 22] chose to work with WOS.

On the other hand, VOSviewer is the most common data analysis tool used in all the studies we compared. Djeki [22] extracted all the results of his research using only VOSviewer. Others like [Das, 21] and [Azurduy, 21] used VOSviewer and Excel, which were also adapted in our research.

In addition to the aforementioned tools, [Gao, 22] used Biblioshiny, and [Morales Muñoz, 22] worked with Citesspace and Knowledgematrix plus. Our bibliometric study is the only one that adapted Python to analyse the results, along with VOSviewer and Excel.

Our study covers the longest period (66 years) compared to other papers, and was richer in results. In order to conduct an in depth comparison, some common results were chosen, such as the most productive year, most prolific countries, resources, institutions, and most cited publications.

The results of the top year regarding the number of publications were similar in all the studies. The years 2019/2020/2021 were the trend years concerning research about E-learning. This can be due to the propagation of COVID-19 and the great interest that distance learning has known since this period.

The most prolific countries were the USA and the UK in almost all studies. The same result of the most prolific sources as [Das, 21] were obtained, which is “Lecture note in computer science”. The other studies obtained different results. Djeki [22] found that “Computers in Human Behavior” and “Computers and Education” are the most

prolific sources. However, the analysis of [Gao, 22] revealed that "International Journal Of Emerging Technologies In Learning" and "Computer And Education" are the most prolific sources. Morales Muñoz [22] and Azurduy [21] did not address this point in their study.

The findings of the most prolific institutions were completely different. Our study showed that the University of California is the most prolific institution, while the University of Politehn Bucuresti was the result of the research of [Djeki, 22]. According to [Das, 21], Kansas State University is the most productive institution, whilst [Morales Muñoz, 22] stated that the University of Toronto is the institution with the highest number of publications.

The most cited papers were also different in each study. Our research reveals that "Empirical Examination of the Adoption of WebCT Using TAM" is the most cited publication. "Investigating Users' Perspectives on E-Learning: an Integration of TAM and IS Success Model" was the most cited publication according to [Djeki, 22], while [Gao, 22] found that "Mobile Learning: a Framework and Evaluation" was the paper with the highest number of citations. Finally the most cited publications according to both [Gao, 22] and [Azurduy, 21], were respectively, "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization" and "E-Learning Success Determinants: Brazilian Empirical Study".

Bibliometric studies	Our Study	Paper1 [Djeki, 22]	Paper2 [Gao ,22]	Paper3 [Das, 21]	Paper4 [Azurduy, 21]	Paper5 [Morales Muñoz , 22]
Period	1954-2020	2015-2020	≤2020	1970-2020	2006-2021	2020-2021
Database	Scopus	WOS	Scopus	Scopus	Lent	WOS
Number of Publication	10,603	12,272	1985	27979	414	602
Data Analyse Tools	Excel Python Vosviewer	Vosviewer	Biblioshiny And Vosviewer	VOS Viewer, And MS-Excel	Excel And Vosviewer	Vosviewer, Citespace, And Knowledgematrix Plus
Top Year Publications	834 on 2019	2265 on 2016	281 on 2020	2185 on 2020	84 on 2020	372 on 2021
Most Prolific Countries	USA, China, UK.	Spain, USA.	USA, UK.	USA, UK.	Malysia, USA, UK.	USA, UK.
Most Prolific Resources	Lecture Notes in Computer Science; Computer and	Computers in Human Behaviour; Computer	International Journal of Emerging Technologies in Learning;	Lecture Notes in Computer Science, ACM; International	-	-

	Education	s and Education	Computer and Education	Conference Proceeding Series		
Most Prolific Institution	University of California	University of Politehn Bucuresti	-	Kansas State University, United States	-	University of Toronto
The Most Cited Publications	Empirical Examination of the Adoption of WebCT Using TAM	Investigating Users' Perspectives on E-Learning: an Integration of TAM and IS Success Model	Mobile Learning: a Framework and Evaluation	Adaptive Sub-gradient Methods for Online Learning and Stochastic Optimization	E-Learning Success Determinants: Brazilian Empirical Study	-

Table 19: Comparative study

4.8 Limitations

Despite considerable efforts to improve the validity of the bibliometric analysis results, the analysis carried out above has several limitations. Undoubtedly, having limited most of our bibliometric analysis to a single bibliometric database is the most obvious limitation of this study, mainly Scopus which is generally regarded as the best one for representing academic research [Valenzuela, 17]. However, the large search string chosen containing the different spellings of the term "e-learning" and "software engineering" for example "elearning", "e-learning", "distance", "virtual " among others , makes it possible to select the largest number of publications in the studied subject. The present bibliometric study covers a long period of 66 years; in addition, a variety of results has been included which allows to have a global idea on the research in this subject. In future research, other databases can be gathered to collect data and reveal similarities and differences among research studies. Besides, the 2020 latest publications that were accepted but not published yet were ignored. Nevertheless, such restrictions are expected to have a minor impact on the findings of this study.

5 Conclusions

The development of E-learning products and the provision of E-learning opportunities is one of the most rapidly expanding areas of education and training [Nariman, 21]. Modern corporations, institutions, universities and schools consider E-learning as a way of educating larger groups of students in less time and reducing the use of various resources [Abdulmajeed, 21] [López-Jiménez et al. 2022]. In particular, the software engineering education system must provide students with the needed resources to

achieve the appropriate educational goals. Researchers have worked to develop and establish relevant learning methods in the field to combine technology with practical learning activities and produce an innovative learning environment that enhances the study. This area has received great interests, especially after the spread of COVID-19. However, at the level of bibliometric studies, no study has been proposed in the subject. Bibliometric studies are very useful to obtain a general picture of the most significant issues occurring in a specific field [Song, 19]. Additionally, the results of any bibliometric study may also provide clear research directions to be developed in the future based on the current mainstreams of a given research field [Cortés-Sánchez, 20]. The present study was aimed to present a bibliometric analysis of the publications on the E-learning Software Engineering education by using several bibliometric indicators and providing some insightful results generated from the Scopus database. Various descriptive analyses, such as the temporal evolution of the publications, the most prolific authors and institutions, the geographical distribution of the publications, the languages used in the publications and the type of publications most relevant to the topic were analysed and visualized. In addition, the paper analyses the author's production, journal productivity, scientific cooperation, authors' citation analysis and journal citation analysis, among other bibliometric parameters.

The results reveal that the number of scientific publications on E-learning in Software Engineering education continues to rise exponentially between 1954 and 2020. Regarding the most productive research institution the University of California (Berkeley, San Diego and Los Angeles) was found. The top author in the research topic based on total papers during the analysed period and total citations is Mario Piattini. In terms of countries, the findings show that the USA and China are the most productive countries in the field followed by the UK, Germany, Spain, Canada, Australia, Italy and Brazil. The research has concluded that English is the most common language used in publications in the field of E-learning Software Engineering education. The most common keywords found in our study were E-learning, Software Engineering, education and Machine Learning. The results found that Lecture notes in computer science are the most common sources in which it is possible to find publications related to this area.

This study has contributed to the existing knowledge by using bibliometric analysis to empirically extend and complement previous studies on E-learning. The results offer important findings on the trends, contributors, and hotspots in the overall development of the E-learning educational community. E-learning in Software Engineering education has proven itself critical in maintaining the continuity of university teaching and research operations during an unforeseeable crisis like the COVID-19 pandemic. It is therefore imperative for universities to make strategic plans and take actions to build sustainable ways of E-learning for future implementation, particularly with relevant technologies and well-established infrastructure.

The research findings will help interested researchers in the field of E-learning to understand the current state of E-learning Software Engineering Education and discover the different research trends. The findings may be used by E-learning institutions to evaluate the quality as a strategic dimension. This bibliometric study is useful for the future research since it highlights the most influential works produced in the field and helps researchers to provide a foundation for their new research. Finally, it should be noted that E-learning Software Engineering education involving fields such as gamifications [Sobrino-Duque et al. 2022] is a promising area for researchers. In

conclusion, as revealed in the results, "Machine Learning" is commonly found in the keywords. Thus, we could expect that the future research target in the field of E-learning will be on the use of machine learning in solving issues in the E-learning sector. Future bibliometric studies can be done on the domain of Machine Learning applied on the E-learning realm.

References

[Abdulmajeed, 21] B. Abdulmajeed, S. Mohammed, B. Abdulla, and A. Abdullah, "An Empirical Investigation on the Impact of the New Electronic Educational Facilities at Kurdistan Universities as Related to E- education Management System," vol. 2, no. 2, pp. 1–10, 2021.

[Aloia, 19] L. Aloia and A. A. Vaporciyan, "E-Learning Trends and How to Apply Them to Thoracic Surgery Education," Thorac. Surg. Clin., vol. 29, no. 3, pp. 285–290, 2019, doi: 10.1016/j.thorsurg.2019.03.006.

[Amin Almaiah, 20] M. Amin 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, 2020, doi: 10.1007/s10639-020-10219-y.

[Ammon, 01] U. Ammon, The Dominance of English as a Language of Science. 2001. doi: 10.1515/9783110869484.

[Andrés, 09] A. Andrés, Measuring Academic Research: How to Undertake a Bibliometric Study. 2009. doi: 10.1533/9781780630182.

[Aparicio, 14] M. Aparicio, F. Bacao, and T. Oliveira, "Trends in the E-Learning Ecosystem: A Bibliometric Study Research-in-Progress," Proc. 20th Am. Conf. Inf. Syst., no. December 2016, 2014.

[Asgariid, 21] S. Asgariid, J. Trajkovic, M. Rahmani, W. Zhang, R. C. Lo, and A. Sciortino, "An observational study of engineering online education during the COVID-19 pandemic," 2021, doi: 10.1371/journal.pone.0250041.

[Ayu, 18] M. Ayu, "ONLINE LEARNING : LEADING E- LEARNING AT HIGHER," no. 4, 2018.

[Azurduy, 21] O. Azurduy, "A BIBLIOMETRIC ANALYSIS OF THE ADOPTION AND USE OF E-LEARNING IN HIGHER EDUCATION Un análisis bibliométrico sobre el uso y la adopción de la educación en línea en la enseñanza superior," vol. VIII, no. No 2.

[Beaie, 09] P. Beaie, S.T. & Acol, "Population and demographic measures: Concepts and definitions for basic MDG indicators," Kingst. Georg. Guyana Bur. Stat., 2009.

[Bosman, 20] L. Bosman, S. Roy, W. McDonald, and C. Ababei, "Using online discussions to connect theory and practice in core engineering undergraduate courses," Comput. Appl. Eng. Educ., vol. 28, no. 3, pp. 675–691, 2020, doi: 10.1002/cae.22238.

[Chang, 20] C. L. Chang and M. Fang, "E-Learning and Online Instructions of Higher Education during the 2019 Novel Coronavirus Diseases (COVID-19) Epidemic," J. Phys. Conf. Ser., vol. 1574, no. 1, 2020, doi: 10.1088/1742-6596/1574/1/012166.

[Chernev, 22] "29 Astonishing E-learning Statistics for 2022." https://techjury.net/blog/elearning-statistics/#gref (accessed Mar. 07, 2022).

[Cnnic, 17] Cnnic, "Statistical Report on Internet Development in China China Internet Network Information Center," no. January, p. 88, 2013, [Online]. Available: http://www1.cnnic.cn/IDR/ReportDownloads/201209/t20120928_36586.htm

[Cortés-Sánchez, 20] J. D. Cortés-Sánchez, "A bibliometric outlook of the most cited documents in business, management and accounting in Ibero-America," Eur. Res. Manag. Bus. Econ., vol. 26, no. 1, pp. 1–8, 2020, doi: 10.1016/j.iedeen.2019.12.003.

[Das, 21] S. Das and P. Assistant, "Research Trends of E - Learning : A Bibliometric and Visualisation Analysis," no. April, 2021.

[Djeki, 22] E. Djeki, J. Dégila, C. Bondiombouy, and M. H. Alhassan, "E-learning bibliometric analysis from 2015 to 2020," J. Comput. Educ., no. 0123456789, 2022, doi: 10.1007/s40692-021-00218-4.

[Elfaki, 19] N. Elfaki, R. Abdelrahim, N. Khalil Elfaki, I. Abdulraheem, and M. R. Abdulrahim, "Impact of e-learning vs traditional learning on students' performance and attitude The Impact of COVID-19 on teaching anatomy; Faculty and student perspective View project Impact of e-learning vs traditional learning on students' performance and attitude," Artic. Int. Med. J., vol. 24, no. 03, 2019, [Online]. Available: https://www.researchgate.net/publication/338528127

[Emami, 20] E. Emami, "COVID-19: Perspective of a Dean of Dentistry," JDR Clin. Transl. Res., vol. 5, no. 3, pp. 211–213, 2020, doi: 10.1177/2380084420929284.

[Eurostat, 22] "Interest in online education grows in the EU - Products Eurostat News - Eurostat." https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20220124-1 (accessed Mar. 09, 2022).

[Farhat, 20] R. Farhat, Y. Mourali, M. Jemni, and H. Ezzedine, "An overview of machine learning technologies and their use in e-learning," Proc. 2020 Int. Multi-Conference Organ. Knowl. Adv. Technol. OCTA 2020, pp. 8–11, 2020, doi: 10.1109/OCTA49274.2020.9151758.

[Fatima, 19] N. Fatima and K. S. Abu, "E-learning research papers in web of science: A biliometric analysis," Libr. Philos. Pract., vol. 2019, 2019.

[Ferguson, 11] G. Ferguson, C. Pérez-Llantada, and R. Plo, "English as an international language of scientific publication: A study of attitudes," World Englishes, vol. 30, no. 1, pp. 41–59, 2011, doi: 10.1111/j.1467-971X.2010.01656.x.

[Gao, 22] Y. Gao, S. L. Wong, M. N. Khambari, and N. Noordin, "A Bibliometric Analysis of the Scientific Production of e-Learning in Higher Education (1998-2020)," vol. 12, no. 5, 2022, doi: 10.18178/ijiet.2022.12.5.1632.

[García-Berná, 19] J. A. García-Berná et al., "Green IT and sustainable technology development: Bibliometric overview," Sustain. Dev., vol. 27, no. 4, pp. 613–636, 2019, doi: 10.1002/sd.1927.

[Gupta, 17] V. Gupta and N. Jain, "Harnessing information and communication technologies for effective knowledge creation: Shaping the future of education," J. Enterp. Inf. Manag., vol. 30, no. 5, pp. 831–855, 2017, doi: 10.1108/JEIM-10-2016-0173.

[Johnson, 20] N. Johnson, G. Veletsianos, and J. Seaman, "U.S. faculty and administrators' experiences and approaches in the early weeks of the COVID-19 pandemic," Online Learn. J., vol. 24, no. 2, pp. 6–21, 2020, doi: 10.24059/olj.v24i2.2285.

[Houaria, 21] C. Houaria and B. Fatiha, "El-Tadwin journal The Scientific Research and the English Language in a Globalized World Case study: The Scientific Publication Criteria in the Algerian University The Scientific Research and the English Language in a Globalized World Case study: The Scie," no. August, 2021.

[Jokiaho, 18] A. Jokiaho, B. May, M. Specht, and S. Stoyanov, "Barriers to using E Learning in an Advanced Way," Int. J. Adv. Corp. Learn., vol. 11, no. 1, p. 17, 2018, doi: 10.3991/ijac.v11i1.9235.

[Karen, 98] Karen W. Arenson, "More Colleges Plunging Into Uncharted Waters of On-Line Courses - The New York Times," 1998. https://www.nytimes.com/1998/11/02/us/more-colleges-plunging-into-uncharted-waters-of-on-line-courses.html?auth=link-dismiss-google1tap (accessed Mar. 21, 2022).

[Kasani, 19] H. A. Kasani and G. S. Mourkani, "A research synthesis of critical success factors of e-learning: A model development," Technol. Educ. J., vol. 13, no. 1, pp. 25–39, 2019, [Online]. Available: https://jte.sru.ac.ir/article_805_en.html

[Kentnor, 15] H. Kentnor, "远程教育和美国在线学习的进展," vol. 17, no. 1, pp. 22–34, 2015, [Online]. Available: https://digitalcommons.du.edu/cgi/viewcontent.cgi?article=1026&context=law_facpub

[Li, 20] L. Li, M. Lin, X. Wang, P. Bao, and Y. Li, "Preparing and responding to 2019 novel coronavirus with simulation and technology-enhanced learning for healthcare professionals: Challenges and opportunities in China," BMJ Simul. Technol. Enhanc. Learn., vol. 6, no. 4, pp. 196–198, 2020, doi: 10.1136/bmjstel-2020-000609.

[Lin, 19] Y. T. Lin, Impacts of a flipped classroom with a smart learning diagnosis system on students' learning performance, perception, and problem solving ability in a software engineering course, vol. 95. Elsevier B.V., 2019. doi: 10.1016/j.chb.2018.11.036.

[Linnenluecke, 20] M. K. Linnenluecke, M. Marrone, and A. K. Singh, "Conducting systematic literature reviews and bibliometric analyses," Aust. J. Manag., vol. 45, no. 2, pp. 175–194, 2020, doi: 10.1177/0312896219877678.

[Liu, 20] Z. Y. Liu, N. Lomovtseva, and E. Korobeynikova, "Online learning platforms: Reconstructing modern higher education," Int. J. Emerg. Technol. Learn., vol. 15, no. 13, pp. 4–21, 2020, doi: 10.3991/ijet.v15i13.14645.

[López-Jiménez et al. 2022] J. J. López-Jiménez et al., "Taking the pulse of a classroom with a gamified audience response system," Comput. Methods Programs Biomed., vol. 213, 2022, doi: 10.1016/j.cmpb.2021.106459.

[Ma, 16] Y. Ma, M. Dong, K. Zhou, C. Mita, J. Liu, and P. M. Wayne, "Publication trends in acupuncture research: A 20-year bibliometric analysis based on pubMed," PLoS One, vol. 11, no. 12, pp. 1–15, 2016, doi: 10.1371/journal.pone.0168123.

[Mathivanan, 21] S. K. Mathivanan et al., "Adoption of E-Learning during Lockdown in India," Int. J. Syst. Assur. Eng. Manag., 2021, doi: 10.1007/s13198-021-01072-4.

[Morales Muñoz, 22] I. Morales Muñoz et al., "E-Learning Research Trends in Higher Education in Light of COVID-19: A Bibliometric Analysis," Front. Psychol, vol. 12, p. 762819, 2022, doi: 10.3389/fpsyg.2021.762819.

[Moubayed, 18] A. Moubayed, M. Injadat, A. B. Nassif, H. Lutfiyya, and A. Shami, "E-Learning: Challenges and Research Opportunities Using Machine Learning Data Analytics," IEEE Access, vol. 6, pp. 39117–39138, 2018, doi: 10.1109/ACCESS.2018.2851790.

[Nariman, 21] D. Nariman, Impact of the interactive e-learning instructions on effectiveness of a programming course, vol. 1194 AISC. Springer International Publishing, 2021. doi: 10.1007/978-3-030-50454-0_61.

[Newbigcing, 59] P. L. Newbigcing, "Required Reading for Anyone Actively Interested in Perception, But Not Fully Familiar With the Literature on Satiation Theory.," vol. 14, no. 1, pp. 60–61, 1959.

[Palvia, 18] S. Palvia et al., "Online Education: Worldwide Status, Challenges, Trends, and Implications," J. Glob. Inf. Technol. Manag., vol. 21, no. 4, pp. 233–241, 2018, doi: 10.1080/1097198X.2018.1542262.

[Pappano, 12] L. Pappano, "The Year of the MOOC NY Times," New York Times, pp. 1–7, 2012, [Online]. Available: https://eds.a.ebscohost.com/eds/detail/detail?vid=0&sid=0c3abe58-6f3d-4cc6-ab8c-a89e898911fb%40sessionmgr4007&bdata=JnNpdGU9ZWRzLWxpdmU%3D#AN=edsgit.A309203065&db=edsgit%0Ahttp://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-

[Pinto, 17] L. Pinto, J. Davidson, R. Sukthankar, and A. Gupta, "Robust adversarial reinforcement learning," 34th Int. Conf. Mach. Learn. ICML 2017, vol. 6, pp. 4310–4319, 2017.

[Razzaque, 20] A. Razzaque, "M-Learning Improves Knowledge Sharing Over e-Learning Platforms to Build Higher Education Students' Social Capital," SAGE Open, vol. 10, no. 2, 2020, doi: 10.1177/2158244020926575.

[Recke, 21] M. P. Recke and S. Perna, "Emergent narratives in remote learning experiences for project based education," Electron. J. e-Learning, vol. 19, no. 2, pp. 59–70, 2021, doi: 10.34190/ejel.19.2.2142.

[Singh, 19] V. Singh and A. Thurman, "How Many Ways Can We Define Online Learning? A Systematic Literature Review of Definitions of Online Learning (1988-2018)," Am. J. Distance Educ., vol. 33, no. 4, pp. 289–306, 2019, doi: 10.1080/08923647.2019.1663082.

[Sobrino-Duque et al. 2022] R. Sobrino-Duque, N. Martínez-Rojo, J. M. Carrillo-de-Gea, J. J. López-Jiménez, J. Nicolás, and J. L. Fernández-Alemán, "Evaluating a gamification proposal for learning usability heuristics: Heureka," Int. J. Hum. Comput. Stud., vol. 161, p. 102774, 2022.

[Song, 19] Y. Song, X. Chen, T. Hao, Z. Liu, and Z. Lan, "Exploring two decades of research on classroom dialogue by using bibliometric analysis," Comput. Educ., vol. 137, no. April, pp. 12–31, 2019, doi: 10.1016/j.compedu.2019.04.002.

[Soni, 20] V. D. Soni, "Global Impact of E-learning during COVID 19," SSRN Electron. J., no. June, 2020, doi: 10.2139/ssrn.3630073.

[Sónia Rolland, 21] E. Technology, "Two Decades of Research in e-Learning : A Deep Bibliometric Analysis," vol. 11, no. 9, pp. 398–404, 2021, doi: 10.18178/ijiet.2021.11.9.1541.

[Talaat Abumandour, 18] E.-S. Talaat Abumandour, "Applying e-learning system for engineering education-challenges and obstacles", doi: 10.1108/JRIT-06-2021-0048.

[Tibaná-Herrera, 18] G. Tibaná-Herrera, M. T. Fernández-Bajón, and F. de Moya-Anegón, "Global analysis of the E-learning scientific domain: a declining category?," Scientometrics, vol. 114, no. 2, pp. 675–685, 2018, doi: 10.1007/s11192-017-2592-7.

[Torres Kompen, 19] R. Torres Kompen, P. Edirisingha, X. Canaleta, M. Alsina, and J. M. Monguet, "Personal learning Environments based on Web 2.0 services in higher education," Telemat. Informatics, vol. 38, no. August 2018, pp. 194–206, 2019, doi: 10.1016/j.tele.2018.10.003.

[Valenzuela, 17] L. M. Valenzuela, J. M. Merigó, W. J. Johnston, C. Nicolas, and J. F. Jaramillo, "Thirty years of the Journal of Business & Industrial Marketing: a bibliometric analysis," J. Bus. Ind. Mark., vol. 32, no. 1, pp. 1–17, 2017, doi: 10.1108/JBIM-04-2016-0079.

[Wang, 18] Y. Wang, "An overview of e-learning in China : History , challenges and opportunities," 2018, doi: 10.1177/1745499918763421.

[Yu, 20] D. Yu and X. He, "A bibliometric study for DEA applied to energy efficiency: Trends and future challenges," Appl. Energy, vol. 268, no. April, p. 115048, 2020, doi: 10.1016/j.apenergy.2020.115048.