



Participatory Educational Research (PER)
Vol.11(3), pp. 220-243, May 2024
Available online at <http://www.perjournal.com>
ISSN: 2148-6123
<http://dx.doi.org/10.17275/per.24.43.11.3>

Id: 1409971

Data Literacy and Education: A Science Mapping Study

Çağrı DEMİRTAŞ*

Social Studies Department, Adıyaman University, Adıyaman, Türkiye

ORCID: 0000-0002-7268-6545

Article history

Received:
08.01.2024

Received in revised form:
18.03.2023

Accepted:
01.05.2024

Key words:

Data literacy; social studies;
education; information age,
digital.

Nowadays, large amounts of data are produced in every field, and these data have become a strategic asset for individuals, businesses, states, and societies. Data literacy involves the skills of using these data effectively. The first aim of the current study is to determine in which years and in which periods more research has been conducted on education and data literacy, reveal the trends in this field over time, and understand what is at the forefront in education and data literacy by analyzing changes in research trends. The second aim of this study is to understand which topics have a wider impact in the field by examining the content of the most cited studies, and to evaluate the potential for collaboration by analyzing the collaborative tendencies and main topic focuses of authors and organizations. This study adopted a science mapping method to examine the relationship in the field of education and data literacy and reveal a general condition of the research on this subject. Data for this study were collected through the Web of Science (WoS) database. As a result of the analysis, an apparent increase in research on education and data literacy since 2011 was examined, and a trend that peaked in 2020 but rose again in 2022 was identified. It was shown that technology and big data-oriented studies have gained importance in thematic evolution. This analysis provides a valuable resource for understanding the current state in the field of education and data literacy and identifying strategies for potential areas of future development.

Introduction

Data holds critical importance in various fields today, ranging from the business world to educational systems and from science to technology. In the age of information, data has transcended being merely a quantitative tool and has evolved into a strategic asset. Data is more than charts, graphs, and spreadsheets. It is being used in powerful, sometimes nearly invisible, ways to shape how we view the world and our role within it. In a world where everyone is an author online, data use is at the crux of humans' daily lives. Thus, conducting an evaluation by focusing on data becomes an inevitable necessity (Fontichiaro & Oehrli, 2016). Many disciplines integrate the use of data into their training and practice, such as sales forecasting, trend analysis and procurement in business, to determine when to remove a pitcher in sports, whether to run or pass a soccer ball, or where to serve in a tennis match, physicians in healthcare

* Correspondency: cagridemirtass@gmail.com

to diagnose and then determine action plans (Mandinach & Gummer, 2013). In the field of education, data collected on student achievements, teaching methods, and the effectiveness of educational programs shed light on decision-making processes for educators and school administrators. Data analysis serves as a critical tool to understand student performance, improve teaching strategies, and create positive impacts on the education system.

In scientific and technological research, data collection and analysis processes form the foundation for new discoveries and innovations. Researchers enhance their understanding by analyzing data from experimental results, contributing to the accumulation of information.

In conclusion, data is an indispensable element for making strategic and information-based decisions, optimizing processes, and making accurate predictions about the future in today's world. Businesses, educational institutions, scientists, and various other sectors can increase their success and achieve sustainable development by effectively utilizing data. Therefore, evaluations focusing on data contribute to organizations making more informed and strategic decisions.

A significant increase in the amount of data in the world and the rapid development of data processing-based technology have drawn attention to data. Therefore, there has been a greater need to conduct studies on addressing and examining the data and accordingly categorizing, analyzing, combining, making it meaningful, and using the data. Hence, it is of critical importance to examine data, the first step in accessing the data, information, knowledge and wisdom (DIKW) model (Ackoff, 1989), and data literacy.

In the dictionary of the Turkish Language Association (TLA), data is defined as “the main element that is the basis of a research, a discussion, a judgment; muta, data, done” (Turkish Language Association, 2023). In the dictionary of the Merriam Webster’s Collegiate; a) Factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation, b) Information in digital form that can be transmitted or processed, c) Information output by a sensing device or organ that includes both useful and irrelevant or redundant information and must be processed to be meaningful (Merriam Webster’s Collegiate Dictionary, 2024).

Data consist of basic, unprocessed, and usually unfiltered information (Kelley, 2002). It has been stated that data consist of symbols containing properties of objects and events and simple, unrefined, and usually unfiltered things (Amidon, 1997, Ackoff, 1989, Davenport & Prusak, 2000, Kelley, 2002, Liew, 2007).

Data can be defined as entities that either exist spontaneously in nature or result from human influence and processes. Data may be dependent on or independent of individuals, as well as interrelated or mutually independent. By coming together, they can form a larger whole, or by separating, they can create smaller components. Humanity has utilized, for its own purposes, a portion of the entities found in nature, whether they are produced, obtained, discovered, or perceived. Obtaining and utilizing all data in nature for an individual appears impractical under current conditions, and not all data may hold significance for humans. It can be asserted that data becomes meaningful and comprehensible when used in alignment with an individual's purpose (Ülken, 2016).

Data can be expressed as simple, unrefined numbers, symbols, words, recordings of events, calculations, images, observations, numerical or verbal expressions, etc., obtained by individuals through experimentation, observation or by chance (Demirtaş, 2022). Data represent characteristics of not only objects or events but also living things, ideas, emotions, times, places, and many other phenomena in the world. Unless the data concerns you, nothing can happen. However, once it somehow starts to interest you, everything can happen (Gencer & Altun, 2021).

Considering the definitions of data, it can be stated that they are made based on the data characteristics, which makes it important what the data types are. Data types are as important as the available definitions of data. In the literature, data types are characterized in the following way (Jeffery, 2014).

- Structured
- Semi-structured
- Unstructured
- ❖ Static
- ❖ Dynamic
- ❖ Streamed
- Secure/open
- Private/public
- Toll-free/Toll
- Open government data
- Open data
- Big data
- Small data
- Dense data

It can be stated that definitions of data types are not precise and can be further expanded. The literacy dimension is the main part that determines the meaning of the definition and types of data. In terms of etymology, literacy is used in English, whereas in Russian the word "gramotnost" is preferred for literacy. In Persian, an Indo-European language, the word "sevad" is used. In this context, the concept of "okuryazarlık" comes to the fore in Turkish. Literacy means going beyond the text, thinking multidimensionally and considering the text in its context. One of the meanings of reading is to make a meaningful evaluation within oneself. Today, the term 'functional literacy' is used for literacy, which is a term of science that pedagogy is more interested in. Literacy implies a significant degree of awareness. In addition to knowing, literacy also includes showing advanced/advanced/high-level skills and/or abilities in that direction (İnan, 2021). At the global level, the concept of literacy is expanding to include the individual, society and the environment due to the rapid change of the cultural environment and the role of literacy in ensuring individual, social and cultural development. In an interdisciplinary context, literacy can be defined as the ability to define, understand, interpret, create, calculate and communicate using visual, audio and digital materials (Demirtaş, 2022).

Literacy is the ability to use language and images in rich and varied forms to read, write, listen, speak, view, represent, discuss, and think critically about ideas. Literacy enables us to share information and to interact with others. Literacy is an essential tool for personal growth and active participation in a democratic society (Ontario Ministry of Education, 2023).



Literacy is; a process of receiving information and making meaning from it and the ability to identify, understand, interpret, communicate, compute, and create text, images, and sounds. Literacy development is a lifelong learning enterprise beginning at birth that involves many complex concepts and understandings. It includes the capacity to learn to communicate, read, write, think, explore, and solve problems (Newfoundland and Labrador Department of Education, 2018).

Situations such as which data to choose, how to process it, how to use it, etc., become possible with data literacy. Data literacy is defined as “a person's level of understanding of how to find, evaluate, and use data to know how to teach” (Mandinach and Gummer, 2016). According to Qin and D'Ignazio (2010), data literacy focuses on functional skills in data collection, processing, management, evaluation, and use. It also involves the critical evaluation of data and understanding what data means, including how to read graphs and tables appropriately, how to draw correct conclusions from data, and how to understand when data are used in misleading or inappropriate ways (Carlson et al., 2011, Koltay, 2015).

Vahey et al. (2006), data literacy includes the ability to: formulate and answer questions using data as part of evidence-based thinking; use appropriate data, tools, and representations to support this thinking; interpret information from data; develop and evaluate data-based inferences and explanations; and use data to solve real problems and communicate their solutions. Deahl's (2014), view, data literacy is the ability to understand, find, collect, interpret, visualize, and support arguments using quantitative and qualitative data. Vahey et al. (2012), explained, the ability to formulate and answer data-based questions, use appropriate data and tools, and interpret the information obtained from the data.

Data and data literacy have become integral components of our world that has become data-based. Data has become more important than ever with the digitalizing world and developing artificial intelligence technologies. With the increasing amount of data, the importance of data-driven word processing programs (big language models) that open the door to a major transformation in human life is also increasing. Especially for the training of complex artificial intelligence systems such as big language models, sufficient amount and quality of data is needed. Access to high quality and diverse data for training big language models is also central to data processing. This process involves collecting, cleaning, processing and modeling large datasets. The quality, accuracy and variety of the data largely determine the impact of the dataset used to train the model. Therefore, using high quality and accurate data in the data processing process of large language models enables the model to better learn the structures of the language and be more successful in various language tasks. These models are trained on a large body of language data to capture the complexity and diversity of the language. However, data literacy also plays an important role in understanding this data correctly and using it effectively. Data literacy is an important foundation for understanding data sources, analysis, ethical thinking in data, making effective and conscious data-based decisions, and encouraging innovation. Improving data literacy is a critical skill that empowers individuals to navigate this world full of data. Data literacy contributes to personal development, professional success, and social progress. As years, days, and even hours pass, the importance of data and data literacy will continue to shape education, the business world, and society in general.

Based on the DIKW (Ackoff, 1989), the structure, characteristics, definitions and principles of each field constitute the knowledge of that field and when literacy is applied to it, the literacy

of that field such as media, environment, law, geography, etc. emerges. The first stage of the information creation process is data, and the process of obtaining and processing data constitutes data literacy. In short, it can be said that the first stage of each literacy area is data literacy. The correct structuring of data literacy, which is the first cornerstone of the success of the DIKW scheme in each field, is of critical importance.

The rapidly changing technological and scientific environment requires individuals to understand, interpret, and use data effectively. Social studies aims to develop social science research and analysis skills by providing students with a comprehensive background in understanding and interpreting social dynamics. The broad spectrum of social studies examines different aspects of society, including subjects such as human behavior, cross-cultural interactions, economic systems, and political structures. When integrated with data, this discipline ensures that the skill of using scientific methods is strengthened to understand and explain social events. By supporting the analytical approach of social sciences, data provides students with a strong critical thinking and research background. In this regard, data literacy supports these analytical processes of social studies, providing students with the opportunity to adapt to the complexity of the information age and develop their skills in using data consciously.

Literacy has always been an important component of Social Studies education. In recent years, through the promotion of research in critical theory, the meaning of literacy has broadened to encompass all media and forms of communication. In today's Social Studies classrooms, learners are encouraged to examine, compose, and decode spoken, written, and visual texts to aid in their understanding of content and concepts and to better prepare them for full and effective participation in their community. Additionally, the goals of literacy include not only language development, but also critical engagement with text, visuals, and auditory information (Ontario Ministry of Education, 2023).

Ontario Ministry of Education (2023) explained the importance of literacy in the social studies teaching program as follows; Literacy continues to evolve as the world changes and its demands shift and become more complex. Today's students live with technological innovations that previous generations never experienced. They are accustomed to receiving information quickly, and often in a non-linear format, and they may engage in social interactions using a variety of technologies. Cross-curricular and subject-specific literacy skills are essential to students' success in all subjects of the curriculum, and in all areas of their live.

In the Social Studies curriculum, skills such as perceiving time and chronological thinking, evidence-based inquiry and research, perceiving change and continuity, spatial thinking, geographical inquiry, geographical observation and fieldwork, maps, tables, graphs, figures and diagrams, logical reasoning, critical sociological thinking, historical problem analysis and decision making, which are expressed as "field skills" and form the basis of data literacy. In addition, it can be seen that literacies such as digital literacy, financial literacy, visual literacy, information literacy, data literacy, citizenship literacy, cultural literacy, sustainability literacy, art literacy, etc. are directly expressed as skills in the curriculum (Ministry of National Education [MoNE], 2024). While data literacy is included in the curriculum, it can be said that other skills in the Social Studies curriculum are directly related to data literacy.

Due to the large number of related disciplines that are utilized in academic studies on social



studies, the number of prominent concepts is also increasing. This situation brings along some problems. In particular, it causes some topics or concepts to be overlooked in literature reviews in studies on social studies. For this reason, it is critical to select umbrella topics that cover related concepts in studies on social studies.

In this respect, it is necessary to emphasize the importance of the relationship between education and data literacy. Nowadays, education and data literacy are tightly linked to the distinctive characteristics of the period known as the information age. The education system must adapt to the demands of this new age and equip individuals to cope with the dynamics of the information age. Data literacy aims to provide individuals with the skills of effectively navigating, evaluating, and producing information among large data sets, particularly under the influence of digitalization. Education plays a critical role in adapting to this dynamic change and effectively guiding individuals through this abundance of data. In this regard, it is inevitable for the education system to play an essential role in providing students with data literacy skills.

Education is more than just a process in which academic information is transferred, it aims to prepare individuals for various life skills. Data literacy is one of these skills. Nowadays, there is a wide range of information in every sector, from scientific research to economic analysis, from health data to social media use, and the ability to interact with this information is essential for drawing accurate and meaningful conclusions.

Educational institutions should develop various strategies to provide students with data literacy skills. First, curricula can focus on subjects such as data analysis, statistical thinking, and information management. Practical applications can be offered to students, enabling them to understand these data by interacting with data sets they may encounter in daily life. Moreover, it is important to integrate technological tools into educational processes. Data visualization tools can help students understand complex data sets. By providing students with the skills of using these tools effectively, teachers can strengthen their skills in data analysis.

Data literacy not only benefits students at an individual level, but also at a societal level. Information-based societies can display more sustainable and knowledge-oriented development owing to data-literate individuals. At this point, education has the potential to create an impact that spreads throughout society.

As a result, the relationship between data literacy and education is of critical importance for individuals to succeed in today's rapidly changing information environment. The development of various strategies by educational institutions to provide students with these skills will contribute to the strong presence of societies in the information age.

In this respect, the first aim of the current study is to determine in which years and in which periods more research has been conducted on education and data literacy, reveal the trends in this field over time, and understand what is at the forefront in education and data literacy by analyzing changes in research trends. The second aim of this study is to understand which subjects create a wider impact within the field by examining the content of the most cited studies, evaluate the collaboration potentials by analyzing the joint work trends of authors and organizations and their main subject focuses, and evaluate how the field has evolved on a global scale upon cross-country collaboration and the increase or decrease trends of this cooperation.

The present study can form an important basis for better understanding the current situation in the field of education and data literacy and determining which aspects future research can focus on.

Within the scope of the stated purpose; Research on education and data literacy; The number of researchers according to years and researchers, the number of national and international citations they received, the journals in which they were published and the changes in these journals over the years, the concepts used in the research and the themes in which the concepts evolved, the current status of the collaborations established by the researchers in these studies according to universities and countries were examined.

Method

A science mapping method was employed in this study to examine published research on the relationship between education and data literacy. Science mapping or Webometrics is an important research subject in the field of bibliometrics. The analysis of web content, custom mapping based on web pages, is a developing technique in the field of bibliometric analysis. Manual screening of web pages by a researcher for specific information about particular research groups or subjects is a useful and common practice (Morris and Van Der Veer Martens, 2008). Scientific mapping mostly aims to reveal the structural and dynamic aspects of scientific research in the field of bibliometrics (Small, 1973, Braam, Moed and Van Raan, 1991, Noyons, Moed and Luwel, 1999).

The publications examined in the study were accessed through the Web of Science (WoS) database. The WoS database was preferred in this study because it includes more resources in the field of social sciences compared to other databases to access relevant research, presents search results with different features in comparison with other databases, and the RStudio program used works easily in this database.

Search Strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram was followed to access the publications to be examined in the study (Moher et al., 2009). Accordingly, first, the relevant search terms were defined in the search engine, then the education-related categories of WoS were selected, and finally, the titles, abstracts, and full texts of the viewed publications were reviewed one by one, and the publications to be examined in line with the main purpose of the study were reached. Figure 1 displays the PRISMA diagram followed to access relevant research.



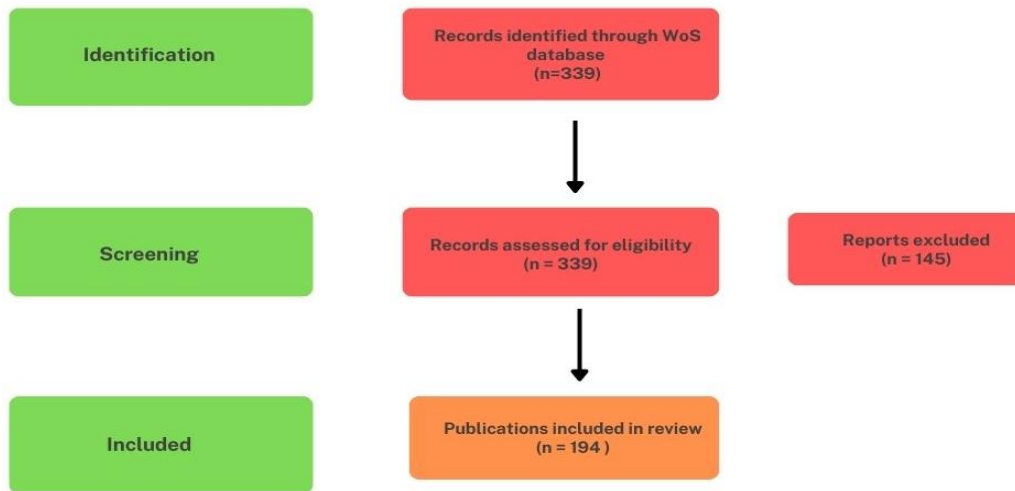


Figure 1. PRISMA Diagram

In this study, while a few direct studies were found in the search for the concept of "social studies" in the Wos database, when the concept of "education" was used, more studies in the content of social studies were found, which constitutes an important example. For this reason, the concept of "education" was chosen instead of "social studies" when designing the study. In this way, the data of the disciplines related to social studies were also accessed.

As seen in Figure 1, to access the publications to be examined in the study, the search terms "education" (All Fields) and "data literacy" (All Fields)" were first defined in the WoS database, and 339 articles were displayed. Articles containing WOS data within the scope of this research. Finally, the researcher examined the titles, abstracts, and full texts of 339 articles one by one, and 145 publications were determined to be unrelated to the study's purpose and were removed from the data set. The 194 publications obtained at the end of this entire process were included in the data analysis. Figure 2 shows the information about the study's data set.

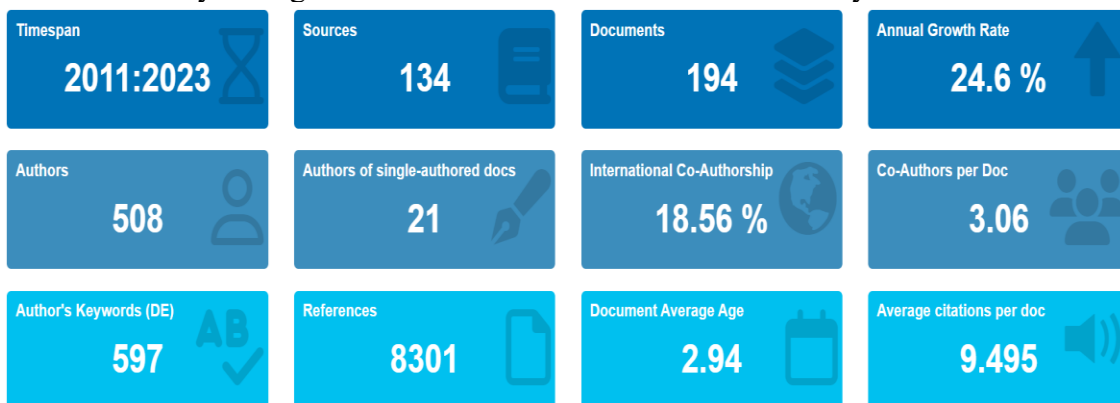


Figure 2. General information about the dataset

As seen in Figure 2, 194 documents from a total of 134 sources were accessed. The publications analyzed in the study were published between 2011 and 2023. The data set obtained contains

information of 508 authors in total. Twenty-one of 194 documents consist of single-authored publications. The data set includes a total of 597 author's keywords and 8301 references. Additionally, it was revealed that the average age of the documents obtained was approximately 3 years, and the average rate of citations per document was 9.49.

The research data were analyzed with the bibliometrix/biblioshiny package of the “R” software (Aria and Cuccurullo, 2017). To this end, the names, author information, abstracts, keywords, and references of 194 publications obtained from the WoS database were combined in a file in the plain text format, and the relevant file was transferred to the RStudio program and analyzed in line with the study's purpose.

Findings

The results obtained by analyzing the data obtained from WoS were evaluated in terms of general information about the research, the relationship between research and the relationships between authors, countries, and institutions where researchers work. A total of 194 studies published between 2011 and 2023 were analyzed in terms of the relationships of the concepts used within the scope of the subject studied. The thematic transformation of the topics investigated in this research has been analyzed. The increase in studies over the years and the citation numbers received by these studies have been examined. Journals in which the scrutinized studies were published and the citation numbers received by these journals within the scope of these studies have also been investigated in this research. The 10 most effective authors, studies, citations, journals, etc., examined in the study were included in the investigation.

Chart 1 displays the distribution of studies on data literacy and education by year.

Annual Scientific Production

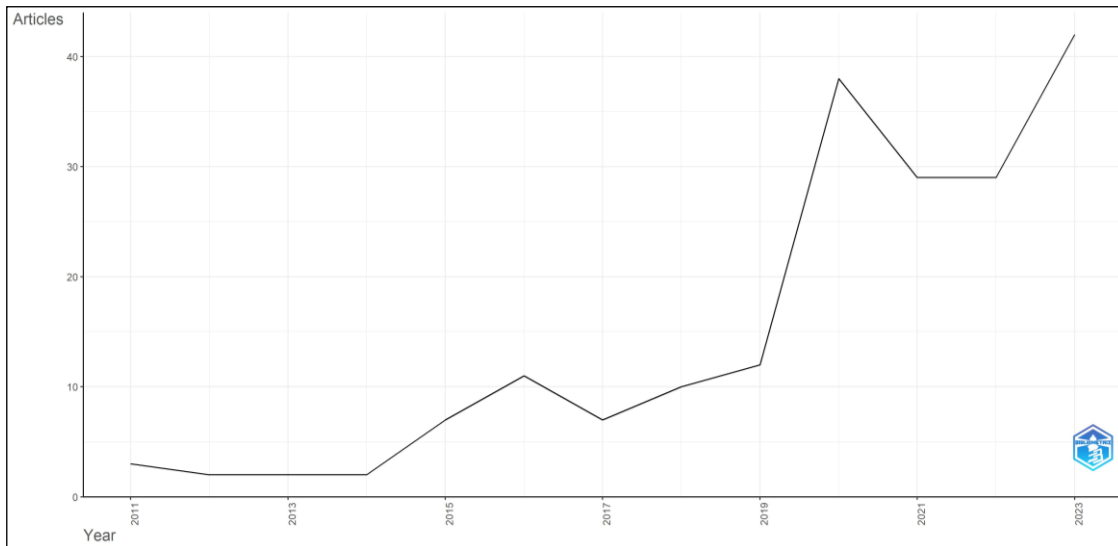


Chart. 1. Annual Scientific Production

As seen in Chart. 1, the first study among publications on digital literacy and education in the



WoS database was accessed in 2011. It can be stated that the number of studies increased after 2014, peaked in 2020, and started to rise in 2023 and reached a new peak. The average number of publications by year was found to be 24.6 between 2011 and 2023.

Average Citation Per Year

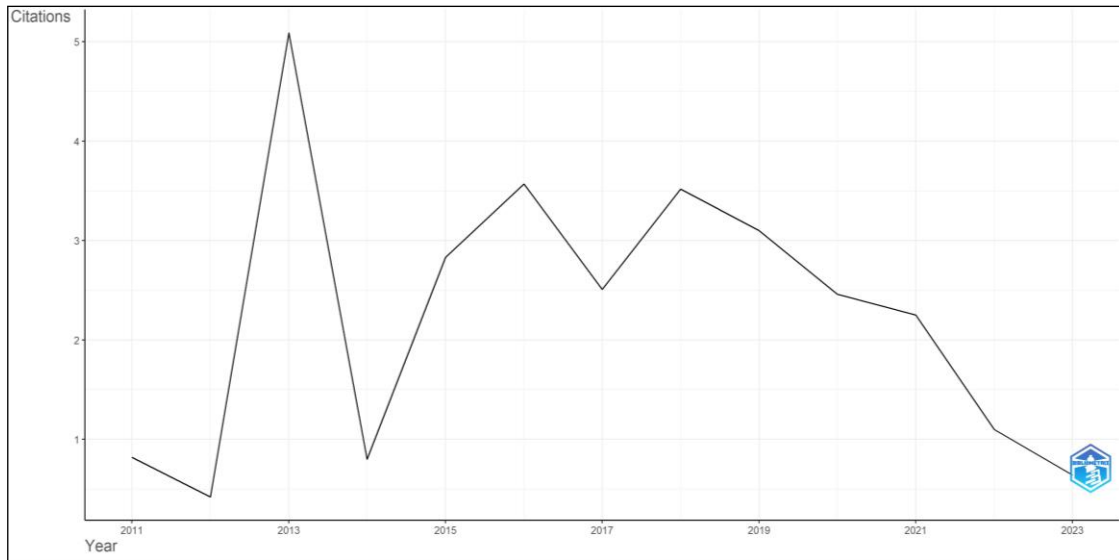


Chart. 2. Average Citation Per Year

Concerning the citations received by publications on data literacy and education in the WoS database, it can be said that an increase started in 2012. This increase peaked in 2013. In the process, a sharp decrease in the citations received by studies and fluctuations in the following period can be observed. It can be noted that the number of citations received by studies decreased continuously after 2018.

Top Ten Most Relevant Sources

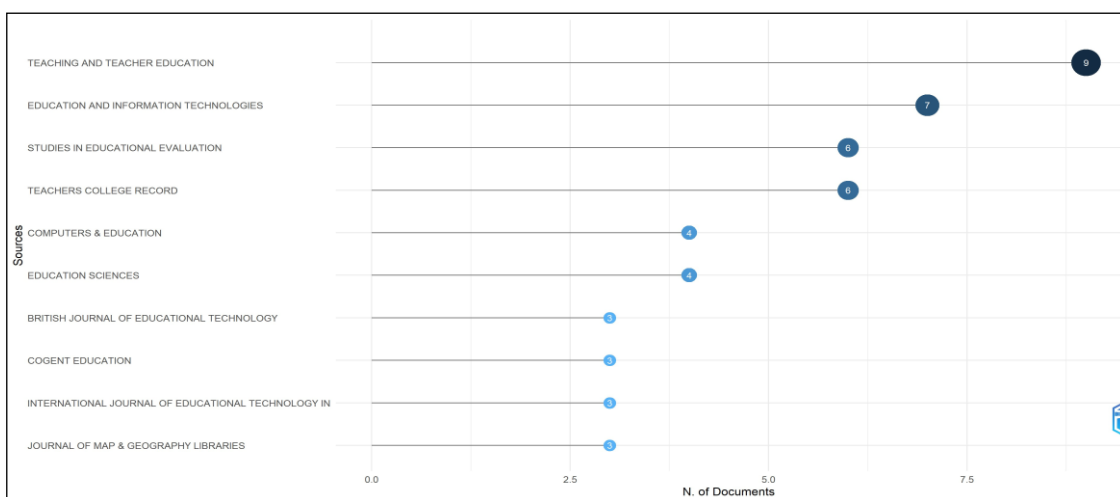


Chart. 3. Most Relevant Sources

The Journal of Teaching and Teacher Education had the highest number of publications on data literacy and education in the WoS database, followed by Education and Information Technologies, Studies in Education Evaluation, and Teachers College Record. The fact that the studies were published in the leading journals in the field is one of the notable points.

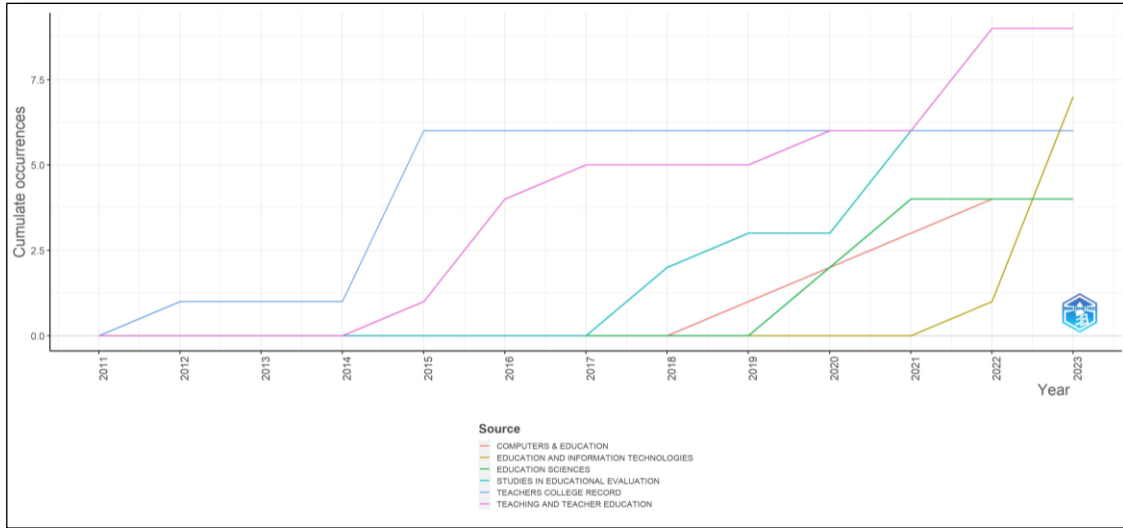


Chart. 4. Source Dynamics

The direction and rate of the increase in the number of publications expressed as Most Relevant Sources can also be observed in the line chart. It can be stated that the Journal of Teaching and Teacher Education displayed a long-lasting increase. The sharpest increase was observed in the Journal of Education and Information Technologies.

Top Ten Most Local Cited Sources

The citation counts of publications related to data literacy and education in the WoS (Web of Science) database have been examined under this heading.

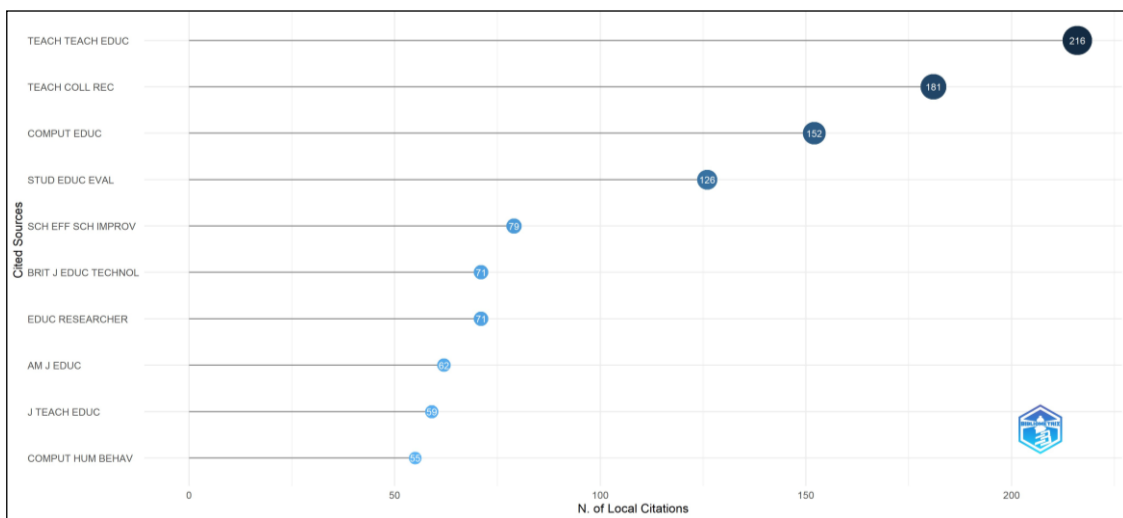


Chart. 5. Most Local Cited Sources



The Journal of Teaching and Teacher Education (216) was the most cited journal on data literacy and education in the WoS database, followed by the journals Teachers College Record (181), Computers & Education (152), and Studies in Education Evaluation (126). The fact that the leading journals of the field were included in the ranking of the most cited journals shows the importance of the subject.

Top Ten Most Influential Authors in Terms of Publications

This study examined the annual number of publications of the authors and their individual contributions to these publications. The data obtained is given graphically.

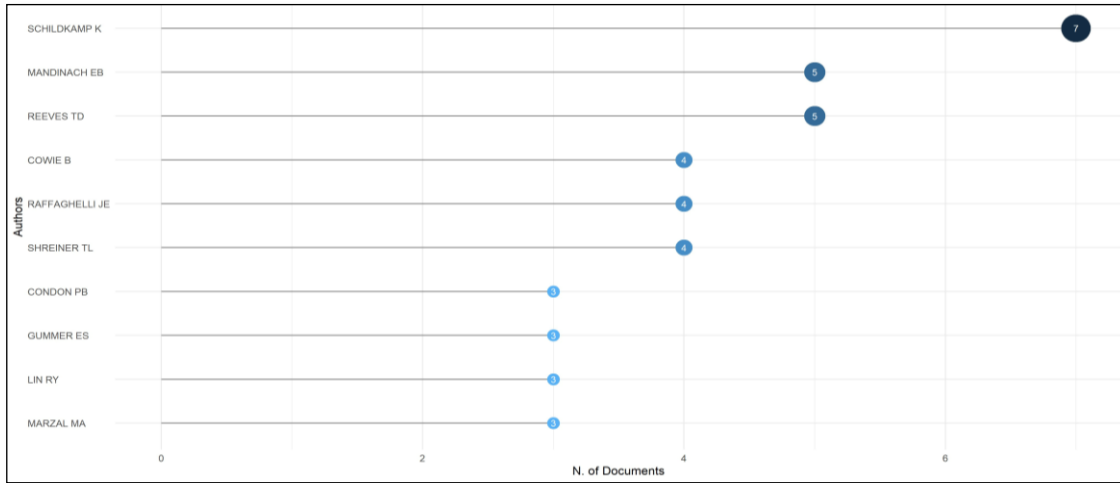


Chart. 6. Most Influential Authors in Terms of Publications

Upon examining the number of publications of the authors in Chart. 6, Schildkamp (TP=7), Mandinach (TP=5), and Reeves (TP=5) come to the fore with their studies. Cowie, Raffaghell, and Shreiner follow the above-mentioned authors with four (4) publications each. Condon, Gummer, Lin, and Marzal follow these authors with three (3) publications each.

Most Relevant Authors

Table. 1 lists the authors’ contributions to the analyzed studies. To this end, the published articles were examined by fractionalized frequency analysis, which is performed by dividing them by the number of authors in each article, as specified by (Aria, Misuraca and Spano, 2020).

Table. 1. Most Relevant Authors

Authors	Articles	Articles Fractionalized
Shreiner, T.L.	4	3.00
Schildkamp, K	7	2.98
Mandinach, E.B.	5	2.33
Reeves, T.D.	5	2.17
Raffaghelli, J.E.	4	2.17
Marzal, M.A.	3	2.00
Cerny, M.	2	2.00



Cowie, B.	4	1.50
Condon, P.B.	3	1.33
Gummer, E.S.	3	1.33

Not: A: Articles, AF: Articles Fractionalized.

Upon examining the data obtained as a result of the analysis, it was found that Shreiner (AF=3.00) was followed by Schildkamp (AF=2.98), Mandinach (AF=2.33), Reeves (AF=2.17), Raffaghelli (AF=2.17), Cerny (AF=2.00), Cowie (AF=1.50), Marzal (AF=2.00), Condon (AF=1.33) and Gummer (AF=1.33).

Top Ten Most Cited Publications

Table. 2 contains the most cited studies on the relationship between data literacy and education in the WoS database and their total citations per year.

Authors -Year-Sources	Article - DOI	TC	TCPY
Mandinach, E.B., Gummer, E. S., 2016, Teach Teach Educ	What does it mean for teachers to be data literate: laying out the skills, knowledge, and dispositions. 10.1016/j.tate.2016.07.011	121	15.13
Prado, J.C., Marzal, M. Á., 2013, Libri	Incorporating data literacy into information literacy programs: core competencies and contents. 10.1515/libri-2013-0010	108	9.82
Schildkamp, K., 2015, Teach Coll Rec	Factors influencing the functioning of data teams. https://doi.org/10.1177/016146811511700403	81	9.00
Hoogland, I., et al., 2016, Teach Teach Educ	Prerequisites for data-based decision making in the classroom: research evidence and practical illustrations. 10.1016/j.tate.2016.07.012	69	8.63
Chen, B.D., et al., 2018, Internet High Educ	Fostering student engagement in online discussion through social learning analytics. 10.1016/j.iheduc.2017.12.002	68	11.33
Schildkamp, K., 2019, Educ Res-Uk	Data-based decision-making for school improvement: research insights and gaps. 10.1080/00131881.2019.1625716	60	12.00
Reisoğlu, I., Çebi, A., 2020, Comput Educ	How can the digital competences of pre-service teachers be developed? Examining a case study through the lens of DigComp and DigCompEdu. 10.1016/j.compedu.2020.103940	53	13.25
López-Meneses, E., et al., 2020, Australas J Educ Tec	University students' digital competence in three areas of the DigCom 2.1 model: a comparative study at three European universities. 10.14742/ajet.5583	50	12.50
Kippers, W.B., et al., 2018, Stud Educ Eval	Data literacy: what do educators learn and struggle with during a data use intervention? 10.1016/j.stueduc.2017.11.001	50	8.33
Mandinach, E.B., Schildkamp, K., 2021, Stud Educ Eval	Misconceptions about data-based decision making in education: an exploration of the literature. 10.1016/j.stueduc.2020.100842	41	13.67

PT= Publication Title, TC= Total Citations, TCPY = Total Citations Per Year

Table. 2. Most Cited Publications

When the two most cited studies are examined, it is seen that the study by Mandinach (2016),



published in the Journal of Teaching and Teacher Education, ranked first with 121 total citations and 15.13 total citations per year. The study by Prado (2013) ranked second with 108 total citations and 9.82 total citations per year. Concerning the content of these studies, the study by Mandinach (2016) addressed a framework that defines the specific knowledge, skills, and attitudes required for teachers to use data in an effective and responsible way (Mandinach & Gummer, 2016). The study by Prado (2013) indicated that it did not aim to contribute to the advancement of data literacy with the proposal of a set of core competencies and contents that can serve as a reference framework for its inclusion in the information literacy programs of libraries (Calzada Prado & Marzal, 2013).

Considering the publications of other authors and their citations, it is seen that the following authors come to the fore: Schildkamp (2015) with 81 total citations and 9 total citations per year, Hoogland (2016) with 69 total citations and 8.63 total citations per year, Chen (2018) with 68 total citations and 11.33 total citations per year, Schildkamp (2019) with 60 total citations and 12 total citations per year, Reisoğlu (2020) with 53 total citations and 13.25 total citations per year, López-Meneses (2020) with 50 total citations and 12.50 total citations per year, Kippers (2018) with 50 total citations and 8.33 total citations per year, and Mandinach (2021) with 41 total citations and 13.67 total citations per year.

Thematic Evolution

Important data were obtained on the thematic evolution of the subjects discussed regarding the relationship between data literacy and education within the process. It can be stated that in the determined evolution, whereas the themes were generally divided and discussed within other themes, new themes emerged as a result of combining several themes and associating them.

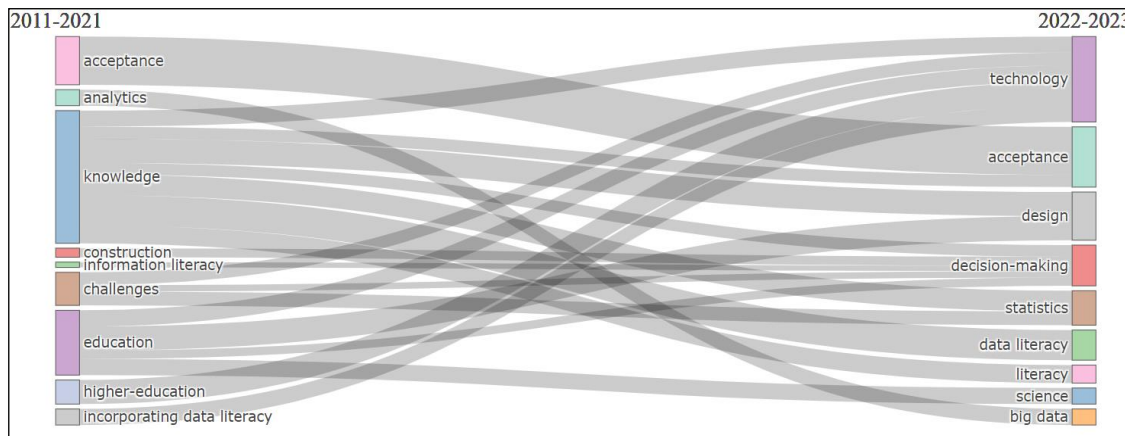


Chart. 7. Thematic Evolution

Upon reviewing the studies, a two-period thematic evolution was identified. The thematic evolution of the studies was divided into the 2011-2021 period and the 2022-2023 period. It was revealed that the themes of acceptance, analytics, knowledge, construction, information literacy, challenges, education, higher-education, and incorporating data literacy emerged between 2011 and 2021 when the studies were published.

The themes of knowledge, education, and higher-education, which emerged between 2011 and

2021 when the studies were published, were divided many times and included in the themes formed between 2022 and 2023. It was found that the themes of technology, acceptance, design, decision-making, statistics, data literacy, science, and big data emerged in the second period of thematic evolution, between 2022 and 2023. It can be said that the theme of technology, which comes to the fore among these themes, is formed by combining the themes of knowledge, education, higher-education, and incorporating data literacy.

Collaboration World Map

In this study, research within the scope of international collaboration on data literacy and education among countries has been examined.

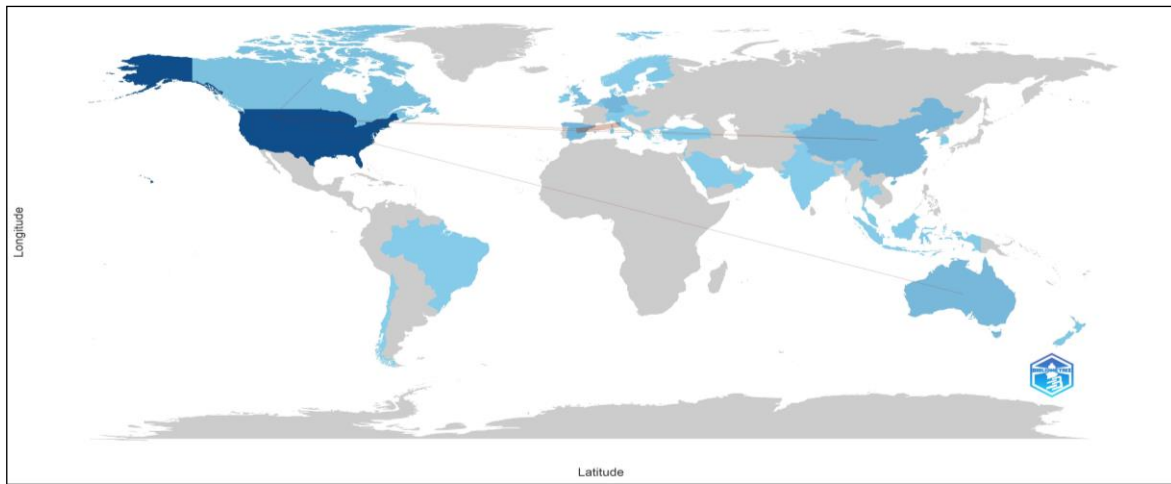


Figure. 3. Collaboration World Map

Upon examining the cross-country collaboration in the studies on data literacy and education at the global level, the network between Spain-Italy, China-Spain, USA-Australia and USA-Canada, USA-China, and USA-Spain emerges. It can be stated that a significant part of the studies were performed in collaboration with Spain. Another important point is that there is collaboration starting from the USA. Considering the countries that participate in collaboration, there are countries on the continents of America, Australia, Asia, and Europe.

Collaboration World Frequency

The frequency data of the studies based on cooperation between countries can be seen in table.3.

Table. 3. Collaboration World Frequency

From	To	Frequency
Spain	Italy	4
China	Spain	2
Usa	Australia	2
Usa	Canada	2
Usa	China	2
Usa	Spain	2

Australia	Germany	1
Australia	Greece	1
Australia	Ireland	1
Australia	Netherlands	1

Countries' Production over Time

This section focuses on the countries where the highest number of studies related to education and data literacy have been conducted.

Table. 4. Countries' Production over Time

Region	Freq
USA	611
Netherlands	135
Australia	123
Belgium	89
Germany	74
Spain	74
China	68
UK	65
Canada	61
Malaysia	18

The analysis of studies determined that most studies on education and data literacy were performed in the USA. While the Netherlands ranked second among the countries with the highest number of studies, Australia ranked third.

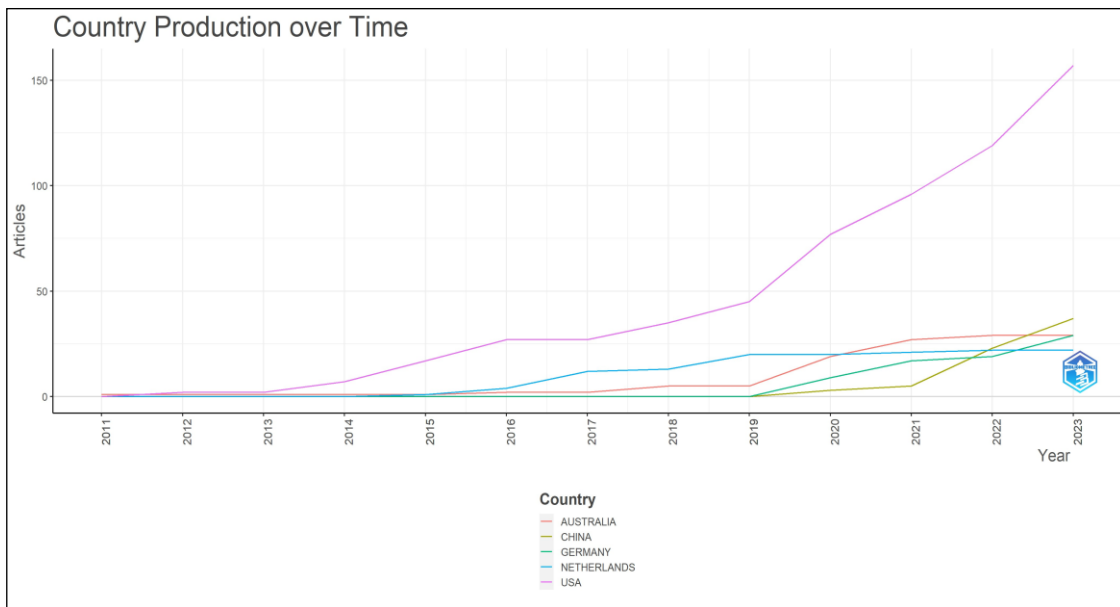


Chart. 8. Country Production Over Time

Upon reviewing studies on a country basis, a continuous and significant increase in studies



conducted in the USA can be observed. It can be seen that there is a similar trend in the studies of other countries.

Most Cited Countries

Considering the scientific study production of the studies, studies conducted in the USA have the highest number of citations. Moreover, a significant difference can be indicated between other countries and the USA. It can be stated that studies produced in the Netherlands and Spain have also reached a significant number of citations.

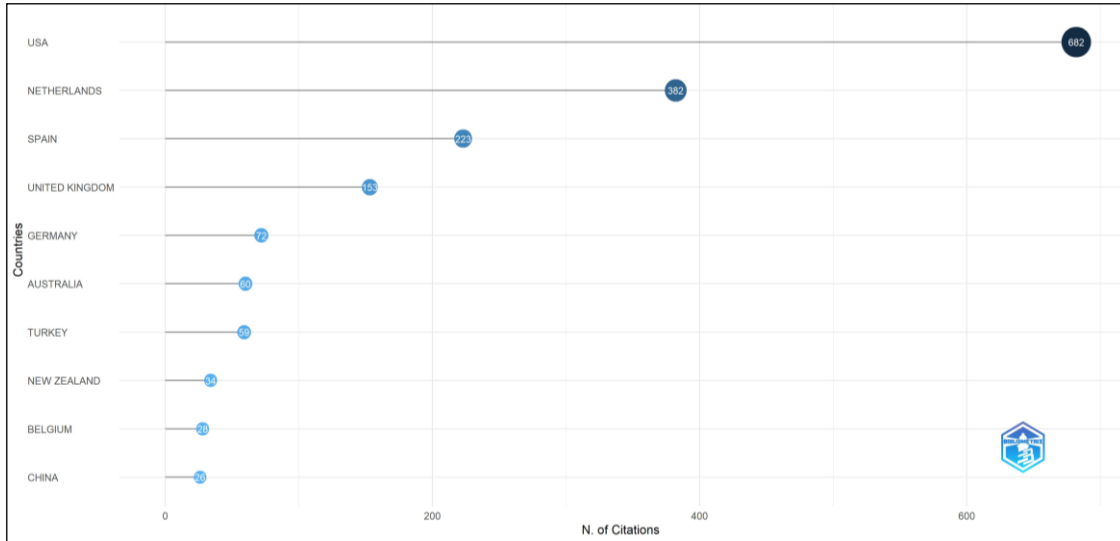


Chart. 9. Most Cited Countries

Most Frequent Words

A point considered important for the study is the frequency of concepts on which the research is built. The data obtained is given graphically.

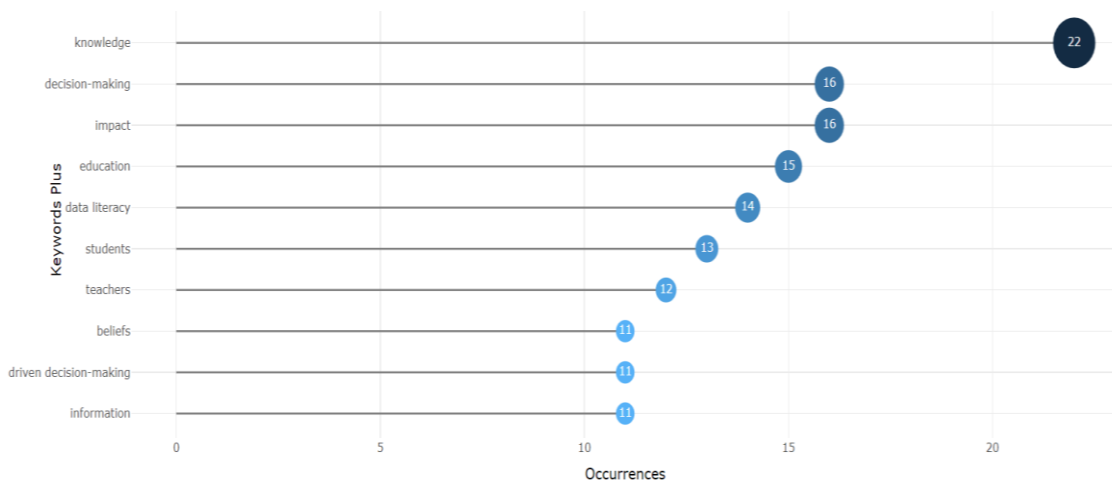


Chart. 10. Frequent Words



The concepts of knowledge (f=22), decision-making (f=16), impact (f=16), education (f=15), data literacy (f=14), students (f=13), teacher (f=12), beliefs (f=11), driven decision-making (f=11), and information (f=11) come to the fore in the research. Concepts are valuable in terms of revealing the basic framework of the studies. It was considered important to evaluate the studies to be conducted within the framework of these concepts. The inclusion of the three concepts that make up the data, information, knowledge and wisdom model [(DIKW), (Ackoff, 1989)] in the chart is another important point.

Word Frequency over Time

The prominent increase in knowledge, one of the concepts examined in the study, can be observed in the line chart. The data obtained is given graphically.

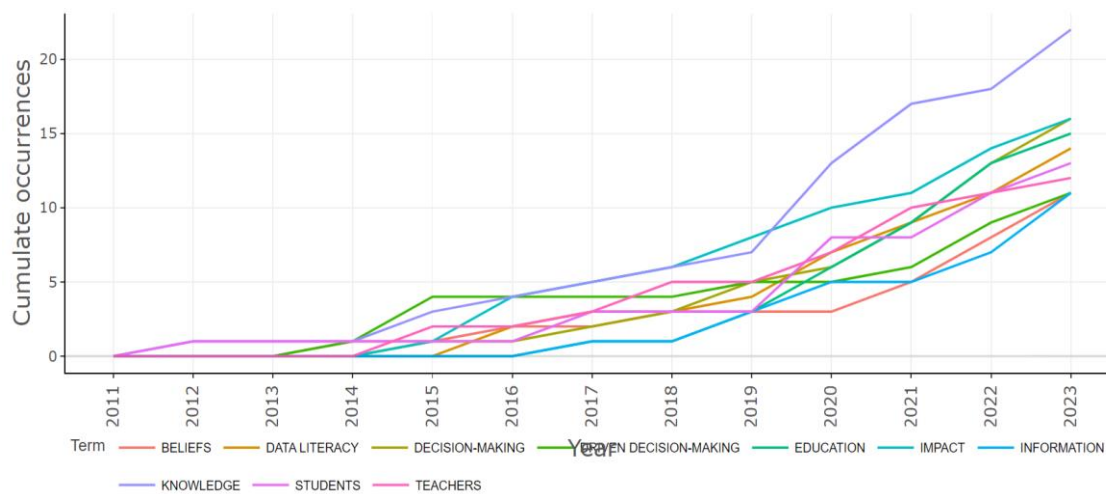


Chart. 11. Word Frequency

Likewise, the concept of impact should also be focused on. Concepts are important in terms of showing the evolution of life. Hence, it is vital to consider conceptual evolution in the construction of studies.

Conclusion and Recommendations

The results of this study were obtained by thoroughly examining the existing information in the field of data literacy and education. Analyses conducted through the Web of Science data encompass the investigation of a total of 194 studies published between 2011 and 2023. Within the overall framework of the research, interactions between these studies, authors, countries, and the institutions where researchers are affiliated have been meticulously evaluated.

The findings reveal significant connections in the realms of data literacy and education. Furthermore, the relationships between the concepts used within the scope of the study have been identified. Through an in-depth examination of the top 10 most influential factors, including authors, studies, citations, journals, etc., the aim was to understand the contributions of these elements to the research field.

In conclusion, this study provides a comprehensive overview of research on the relationship

between data literacy and education, contributing significantly to the literature. The obtained findings can form a foundational framework for future studies in this field and offer guidance to researchers.

The review demonstrates a significant increase in research on education and data literacy, with an increasing trend since 2011. It was observed that this increase peaked in 2020, then there was a decrease, but it started to increase again as of 2022. The aforesaid trend indicates that researchers' interest in these subjects continues and that the field is still actively developing.

However, upon focusing on the number of citations, it was determined that the number of citations peaked in 2013 and displayed a continuous downward trend in the following years. This suggests that the effect of studies that come to the fore in a certain period decreases over time or a certain saturation point is reached. The decrease in the number of citations may provide important clues to future researchers about specific issues they should focus on with regard to the subject.

The analysis of journals reveals that journals such as *Teaching and Teacher Education*, *Education and Information Technologies*, *Studies in Education Evaluation*, and *Teachers College Record* are the leading journals in the field. The fact that the *Journal of Teaching and Teacher Education* included the most studies and received the highest number of citations demonstrates that this journal plays a critical role in the center of the field. It can be said that studies published in this journal contribute significantly to the literature on education and data literacy.

Conceptual analysis shows that concepts such as knowledge, decision-making, impact, education, data literacy, students, teacher, beliefs, driven decision-making, and information play a central role in research on education and data literacy. The above-mentioned concepts represent the superordinate themes that researchers focus on in this field. In this regard, it can be stated that the research conducted around the concepts in question forms the basis of the field and creates a discipline shaped around these concepts.

Thematic evolution analysis demonstrates that the themes of acceptance, analytics, knowledge, construction, information literacy, challenges, education, higher-education, and incorporating data literacy came to the fore during the 2011-2021 period, and the themes of technology, acceptance, design, decision-making, statistics, data literacy, science, and big data formed the superordinate theme between 2022 and 2023. This change refers to an evolution in which technology and big data-oriented studies gain importance.

The analysis of countries and collaborations shows that large-scale collaborations initiated by Spain-Italy and the USA come to the fore, in addition to studies conducted under the leadership of the USA. These collaborations show that global collaborations in the field of education and data literacy are essential and that the field has diverse perspectives.

In the literature review on education and data literacy, although bibliometric studies were not found, regarding the importance of data literacy in education; Shreiner & Guzdial (2022), in a study designed to help teachers provide technology-supported data literacy instruction in social studies, stated that social studies teachers found data visualizations useful and useful for using data visualizations as part of standard and inquiry-based social studies education, teaching



critical analysis of data visualizations, and teaching critical analysis of data visualizations. In the conclusion, it was stated that very few of the teachers participating in the study had not yet adapted the resources provided to their own classrooms. It was also stated that most of the teachers did not teach data literacy as a part of social studies. In a similar study on young adults' understanding and use of data, Gebre (2018) found that students' understanding of data was limited to the contexts of experiments and surveys, utility and usage information, and numerical charts and graphs. However, it was emphasized that students used a wide range of data in their original projects compared to their written explanations. It was stated that the design of data literacy interventions contributed to the students developing a broader awareness of the nature of data in everyday life. Rowe, Riggio, De Amicis & Rowe, (2020), in a study to determine primary and secondary school (K-12) teachers' perceptions of cross-reality (XR) tools for data visualization and the use of sensor data from the built environment in the classroom curriculum, stated that teachers showed relatively high interest in learning about sensor applications and innovative interactive techniques, especially secondary school teachers saw more value in using these applications for teaching and learning. Mahmud & Wong (2022), in a study investigating the perspectives of undergraduate students in Malaysia on the importance of 21st century skills for career preparation, emphasized that due to the demands of an ever-changing world, many institutions have begun to incorporate 21st century skills into curriculum design to better prepare students for success in the workplace and lifelong career development. In this context, the study argues that data literacy is a necessary skill for success in the workplace and similarly, problem solving skills help to develop critical thinking skills that contribute to the development of creative thinking skills.

As a result, this analysis serves as a crucial guide to understanding the current state and potential areas of development in the fields of education and data literacy. It provides valuable insights for researchers, educators, and policymakers seeking to formulate advanced strategies in the realms of education and data literacy. Focusing on increased collaboration and interdisciplinary research is crucial to ensuring the sustainability and amplification of the impact of studies in this field.

This examination offers a detailed understanding of the current state of education and data literacy, shedding light on essential aspects that can guide future research. The findings presented in the study have the potential to guide researchers, educators, and policymakers interested in defining strategies for the future. Emphasizing collaboration and interdisciplinary research is crucial for sustaining and enhancing the impact of efforts in this field.

Future researchers can further enrich the evolution of education and data literacy by concentrating on identified key journals and concepts. Additionally, strengthening international collaborations contributes to diversifying the field globally and introducing new perspectives. In this context, it is anticipated that future studies in education and data literacy will offer a more comprehensive understanding and broader perspective on the subject.

Recommendations

In light of the findings of this bibliometric analysis, we can make a few recommendations to researchers who want to conduct more in-depth research in the field of education and data literacy:

- To understand the reasons for the decrease in the number of citations, a detailed examination of the studies after 2013 and the interactions of these studies may reveal the factors underlying the decrease in the number of citations.
- Focusing on articles published in the leading journals such as Teaching and Teacher Education, Education and Information Technologies may help researchers better understand the future trends of the field.
- Focusing on conceptual studies on key concepts such as knowledge, decision-making, impact, education, and data literacy may be important for understanding the evolution of these terms within the field.
- Conducting a detailed study examining the similarities and differences between the periods 2011-2021 and 2022-2023 to better understand the thematic evolution in these periods may contribute to understanding the reasons for the trends and the effects of these changes on the field's development.
- Analyzing prominent cross-country collaborations such as the USA and Spain-Italy may contribute to understanding the dynamics of collaborations at the global level.

The recommendations presented above can serve as a basis for researchers in the field of education and data literacy to deepen existing findings, shed more light on the development of the field, and lead to new discoveries on this subject.

References

- Ackoff, R. L. (1989). "From Data to Wisdom," *Journal of Applied Systems Analysis*, 16, 3-9.
- Amidon, D. M. (1997). *Innovation Strategy for the Knowledge Economy: The Ken Awakening*. Butterworth-Heinemann.
- Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975.
- Aria, M., Misuraca, M., & Spano, M. (2020). Mapping the evolution of social research and data science on 30 years of social indicators research. *Social Indicators Research*, 1–29. <https://doi.org/10.1007/s11205-020-02281-3>
- Braam, R. R., Moed, H. F., & Van Raan, A. F. (1991). Mapping of science by combined co-citation and word analysis. I. Structural aspects. *Journal of the American Society for Information Science*, 42(4), 233-251.
- Calzada Prado, J., & Marzal, M. A. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63(2), 123-134.
- Canada, Council of Ministers of Education. (2023). *The Ontario curriculum: Social studies, grades 1 to 6; History and geography, grades 7 and 8*. Retrieved from <https://www.dcp.edu.gov.on.ca/en/curriculum/elementary-sshg/grades/g2-ss/downloads>
- Canada, Newfoundland and Labrador Department of Education. (2018). *Social studies 1201 curriculum guide*. Retrieved from https://www.gov.nl.ca/education/files/Social_Studies_1202_2018.pdf
- Carlson, J., Fosmire, M., Miller, C. C., & Nelson, M. S. (2011). Determining data information literacy needs: A study of students and research faculty. *Portal: Libraries and the Academy*, 11(2), 629-657.



- Chen, B., Chang, Y. H., Ouyang, F., & Zhou, W. (2018). Fostering student engagement in online discussion through social learning analytics. *The Internet and Higher Education*, 37, 21-30.
- Davenport, H. T., & Prusak, L. (2000). *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press.
- Deahl, E. (2014). *Better the data you know: Developing youth data literacy in schools and informal learning environments* [Master's thesis, Massachusetts Institute of Technology].
- Demirtaş, Ç. (2022). *Sosyal bilgiler eğitimi için bilgi okuryazarlığı modeli önerisi [A model proposal for knowledge literacy in social studies education]*. Bolu Abant İzzet Baysal University, Graduate School of Education.
- Fontichiaro, K., & Oehrli, J. A. (2016). Why data literacy matters. *Knowledge Quest*, 44(5), 21-27.
- Gebre, E. H. (2018). Young adults' understanding and use of data: Insights for fostering secondary school students' data literacy. *Canadian Journal of Science, Mathematics and Technology Education*, 18(4), 330-341.
- Gençer, R., & Altun, A. (2021). Dijitalleşme, bilgi hiyerarşisini değiştirdi mi? (VEBB: veri, enformasyon, bilgi ve bilgelik) [Has digitalization changed the hierarchy of knowledge (DIKW: data, information, knowledge, and wisdom)?]. *Diyalektolog-Uluslararası Sosyal Bilimler Dergisi [Dialectologist-International Journal of Social Sciences]*, (27).
- Hoogland, I., Schildkamp, K., Van der Kleij, F., Heitink, M., Kippers, W., Veldkamp, B., & Dijkstra, A. M. (2016). Prerequisites for data-based decision making in the classroom: Research evidence and practical illustrations. *Teaching and teacher education*, 60, 377-386.
- İnan, S. (2021). *Siyaset Okuryazarlığı "Yöneten Birey Olmak ve Okullarda Siyaset Eğitimi Mümkün mü?" [Political literacy: Is it possible to be a governing individual and to teach politics in schools?]*. Yeni İnsan Yayınevi. İstanbul.
- Jeffery, K. (2014). *Data is the new oil: Best practices for data management & sharing*. The Joint Research Centre (JRC).
- Kelley, J. (2002). *Knowledge nirvana – achieving the competitive advantage through enterprise content management and optimizing team collaboration*. Xulon Press.
- Kippers, W. B., Poortman, C. L., Schildkamp, K., & Visscher, A. J. (2018). Data literacy: What do educators learn and struggle with during a data use intervention? *Studies in Educational Evaluation*, 56, 21-31.
- Koltay, T. (2015). Data literacy: In search of a name and identity. *Journal of Documentation*, 71(2), 401-415.
- Liew, A. (2007). Understanding data, information, knowledge and their inter-relationships. *Journal of knowledge management practice*, 8(2), 1-16.
- López-Meneses, E., Sirignano, F. M., Vázquez-Cano, E., & Ramírez-Hurtado, J. M. (2020). University students' digital competence in three areas of the DigCom 2.1 model: A comparative study at three European universities. *Australasian Journal of Educational Technology*, 36(3), 69-88.
- Mahmud, M. M., & Wong, S. F. (2022, November). Digital age: The importance of 21st century skills among the undergraduates. In *Frontiers in Education* (Vol. 7, Article 950553). Frontiers Media SA.
- Mandinach, E. B., & Gummer, E. S. (2013). Building educators' data literacy: Differing perspectives. *The Journal of Educational Research & Policy Studies*, 13(2), 1-5.

- Mandinach, E. B., & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366-376.
- Mandinach, E. B., & Schildkamp, K. (2021). Misconceptions about data-based decision making in education: An exploration of the literature. *Studies in Educational Evaluation*, 69, 100842.
- Merriam-Webster. (2024). Data. In *Merriam-Webster's Collegiate Dictionary*. Erişildiği tarih: 4 Mayıs 2024. <https://www.merriam-webster.com/dictionary/data>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*, T. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), 264-269.
- Morris, S. & Van Der Veer Martens, B. (2008). Mapping research specialties. *Annual Review of Information Science and Technology*, 42(1), 213–295.
- Noyons, E. C. M., Moed, H. F., & Luwel, M. (1999). Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study. *Journal of the American society for Information Science*, 50(2), 115-131.
- Qin, J., ve D'Ignazio, J. (2010). Lessons learned from a two-year experience in science data literacy education. In *Proceedings of the 31st Annual IATUL Conference* (pp. 1-10), 20–24 June 2010. Retrieved from <http://docs.lib.purdue.edu/iatul2010/conf/day2/5>
- Reisoğlu, İ., & Çebi, A. (2020). How can the digital competences of pre-service teachers be developed? Examining a case study through the lens of DigComp and DigCompEdu. *Computers & Education*, 156, 103940.
- Rowe, S., Riggio, M., De Amicis, R., & Rowe, S. R. (2020). Teacher perceptions of training and pedagogical value of cross-reality and sensor data from smart buildings. *Education Sciences*, 10(9), 234.
- Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational research*, 61(3), 257-273.
- Shreiner, T. L., & Guzdial, M. (2022). The information won't just sink in: Helping teachers provide technology-assisted data literacy instruction in social studies. *British Journal of Educational Technology*, 53(5), 1134-1158.
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for information Science*, 24(4), 265-269.
- Turkish Language Association. (2024). Veri. *Turkish Language Association Dictionary*. Retrieved May 4, 2024, from <https://sozluk.gov.tr/>
- Türkiye, Ministry of National Education. (2024). *Common text of teaching programs: Social studies course curriculum*. Retrieved from <https://gorusoneri.meb.gov.tr/>
- Ülken, H. Z. (2016). *Bilgi ve değer. [Knowledge and value]* Doğu Batı Yayınları.
- Vahey, P., Rafanan, K., Patton, C., Swan, K., van't Hooft, M., Kratoski, A., & Stanford, T. (2012). A cross-disciplinary approach to teaching data literacy and proportionality. *Educational Studies in Mathematics*, 81, 179-205.
- Vahey, P., Yarnall, L., Patton, C., Zalles, D., & Swan, K. (2006, April). Mathematizing middle school: Results from a cross-disciplinary study of data literacy. In *Proceedings of the Annual Meeting of the American Educational Research Association*, San Francisco, CA (pp. 1-15).

Web of Science. (n.d.). [Summary of relevance marked records on education and data literacy]. Retrieved July 10, 2023, from <https://www.webofscience.com/wos/woscc/summary/marked/relevance/1?markedListId=education%20and%20data%20literacy>