Media and Communication
2025 • Volume 13 • Article 9237
https://doi.org/10.17645/mac.9237

cogitatio

REVIEW

Open Access Journal

Beyond Numeracy, a Data Literacy Topical Scoping Review (2011–2023)

Lotte Vermeire [1], Wendy Van den Broeck [1], Fazlyn Petersen [2], and Leo Van Audenhove [1,3]

[1] imec-SMIT, Vrije Universiteit Brussel, Belgium
[2] Department of Information Systems, University of the Western Cape, South Africa
[3] CoLab for eInclusion and Social Innovation, University of the Western Cape, South Africa

Correspondence: Lotte Vermeire (lotte.vermeire@vub.be)

Submitted: 31 August 2024 **Accepted:** 5 November 2024 **Published:** 23 January 2025

Issue: This review is part of the issue "Evaluating and Enhancing Media Literacy and Digital Skills" edited by Leen d'Haenens (KU Leuven) and Willem Joris (Vrije Universiteit Brussel), fully open access at https://doi.org/10.17645/mac.i466

Abstract

As society increasingly recognises the value of data, proficiency in using and understanding data has become crucial. However, there is no universal consensus on the definition of data literacy. Therefore, this study provides the first extensive, mixed methods scoping review of the topical evolution of data literacy within social and educational sciences from 2011 to 2023. By identifying key themes and research trends, this review offers a comprehensive understanding of the dynamic nature of data literacy. Our sample consists of 210 English-language, peer-reviewed articles from Scopus and Web of Science. The findings reveal a field that is evolving alongside media and digital literacy discussions, with notable growth in publications, particularly in 2019, 2020, and 2023, thus highlighting data literacy's recognition as a distinct paradigm. Data literacy is shifting beyond traditional frameworks, with increasing attention to issues of equity and accessibility—areas still underexplored in current literature. Notably, the research demonstrates a shift from simply developing individual data skills to fostering a socially aware form of data literacy that empowers citizens to critically engage with data and navigate a datafied society actively and responsibly. This review emphasises the need for a nuanced, context-specific approach to data literacy, much like digital literacy, as different demographics and contexts encounter varying needs and challenges. As a dynamic, ever-evolving concept, future research and programs must address these diverse levels of engagement and expertise, ensuring that data literacy is inclusive, adaptable, and supported by social structures.

Keywords

data literacy; data skills; media education; topical evolution; understanding data; using data

1. Introduction

There is an increasing importance of data literacy in today's datafied society (Bhargava et al., 2015; Gray et al., 2018). Data is used across various fields and areas of society, such as in business, policy, and education, and employers are increasingly seeking data-literate employees (Ghodoosi et al., 2023). As a result, data literacy is a concept that has gained traction over the years and is discussed across different disciplines and contexts (Wolff et al., 2016). However, as with the discourse on media literacy (van Dijk, 2020), data literacy has been going through a conceptual evolution, with authors not agreeing on a unified definition or framework. This evolution goes hand in hand with societal changes, such as the emergence of big data and AI, and various disciplines, such as those within social sciences, are starting to pay more attention to the concept. Over the years, various aspects have been emphasised within the definition of data literacy. Earlier definitions mainly focus on data analysis skills (Schield, 2004), while more recent definitions include the competences to understand the individual implications of data (Seymoens et al., 2020). Other related concepts also emerged that were sometimes used interchangeably, such as information literacy (Carlson et al., 2011; Schield, 2004). In their meta-review, Wuyckens et al. (2022) find that there is a conceptual confusion and lack of clarity around media literacy, information literacy, and digital literacy.

On the conceptual ambiguity surrounding different types of digital literacies, Mensonides et al. (2024) traced the historical development of media, information, and digital literacies. They emphasise the need for digital literacies to be flexible and adaptable to the challenges posed by new technologies and to be recontextualised. The authors argue for a shift away from one-size-fits-all understandings of digital literacies, proposing instead that these concepts should be understood as socially situated pedagogical processes able to change and evolve and be influenced by societal changes and contexts (Mensonides et al., 2024). Similar to the understanding of information literacy and other forms of literacy (Wuyckens et al., 2022), data literacy has been defined and interpreted in various ways, with no clear consensus on its scope or the specific competences it entails (Bawden, 2001; Mensonides et al., 2024; Van Audenhove et al., 2020). Due to the specific challenges posed by the increasing availability and complexity of data, academic discourse has been calling for a separate definition (Koltay, 2015). There is a need to distinguish it from other literacies and ensure a comprehensive understanding of its meaning.

Therefore, the purpose of the study is to provide contextual understanding of the topical evolution of data literacy in social and education sciences. As society becomes more data-driven, there is an increasing need for data literacy among citizens. And as data increasingly influences the societal, political, and cultural aspects of our lives, we observe a growing need for data literacy that leaves the confines of data science and statistical analytics and encompasses other disciplines (Bhargava et al., 2015; D'Ignazio, 2017; Gray et al., 2018; Van Audenhove et al., 2024). As Williamson et al. (2020) highlights, datafication in education—where data is used to assess and personalise learning and instruction—serves as an example of how data plays an increasingly central role in shaping educational practices. Alongside these developments, concepts like critical and creative data literacy have emerged, offering new interpretations of and pathways to engaging with data in meaningful ways (D'Ignazio, 2017; Louie et al., 2022). In this study, we focus on the topical evolution of data literacy in education and social science disciplines, such as communication, sociology, library and information studies, psychology, education, and pedagogy. We will identify which main topics and themes can be recognised, which topics carry the most research interest, as well as what trends can be identified. We will then provide a mapping of the development of data literacy research, contributing to a

clearer conceptual framework, focusing on the period from 2011 to 2023—starting from the year that relevant articles were published and included in the covered databases. Consequently, the study aims to answer the following research questions:

> RQ1: What is the topical evolution of data literacy in social and education science publications from 2011 to 2023?
>
> RQ2: What thematic trends can be identified in the topical evolution of data literacy (2011–2023)?
>
> RQ3: What main themes and topics are addressed in the publications on data literacy (2011–2023)?

In previous research, we defined three different fields within data literacy research, namely: the social sciences, open data, and STEM/education fields. The open data field, which focuses on making data freely available to enhance participatory democracy but recognises the need for data literacy among citizens; the STEM education field, which integrates data into curricula to teach students how to use and reflect on data, though it often emphasises usage over understanding; and the social sciences field, which examines the broader societal implications of data, promoting critical reflection on data's role in society (Van Audenhove et al., 2024). This is to be expected, as different disciplines and contexts require unique data competences, such as data interpretation for policy decisions, data literacy for fact-checking, data visualisation development, or using learning analytics to improve education. Other categorisations of data literacy are based on how different communities of practice define the concept. A good example is Ahmed et al. (2021), who conducted a study focused on three categories of communities of practice: education, fields and professions, and citizenship. Each community had a different understanding of the notion of data literacy.

Salomão Filho et al. (2023) conducted a systematic and narrative review of 39 articles that explore the educational and socio-political aspects of data literacy (2015–2021). The authors note the interdisciplinary nature of research on data literacy encompasses a wide range of topics and distinct aspects of data literacy, making it challenging to form a cohesive conceptual understanding. While we recognise that the interdisciplinary nature of the concept of data literacy makes it difficult to identify separate domains, we use the fields identified in Van Audenhove et al. (2024) as a conceptual starting point for understanding how data literacy is approached across different fields. Concretely, we focus on data literacy from educational and societal perspectives, with the education dataset focusing on data literacy within structured learning environments, such as curricula and pedagogical practices, while the social science dataset emphasises broader societal implications, including policymaking and civic engagement. This distinction allows us to analyse data literacy from educational and societal perspectives. Moreover, by treating "open data" as an interdisciplinary concept embedded in both fields, rather than as a separate field, we aim to provide a comprehensive review of how data literacy evolves across education and social sciences.

This study is structured as follows: (a) a methodology section outlining the approach used to select and analyse 210 articles, and (b) the results section, which provides a bibliometric analysis of the literature, an exploration of early definitions and the lack of consensus on data literacy, and a look at the thematic trends and concepts in the research. The review concludes with a summary of key findings, implications for future research, and the study's limitations.

2. Methodology

This study maps the topical evolution of the data literacy concept in the social and education sciences (2011–2023), identifying the main themes, evolutions of the concept, and gaps for future research. Since this is a topical analysis of one concept, a scoping review is a fitting methodology as it allows for a comprehensive mapping of the broader literature. For this, we made use of the methodological framework from Arksey and O'Malley (2005), combining this with the PRISMA guide for scoping reviews (Tricco et al., 2018). Following these guidelines, we first developed a research protocol (see the Supplementary File, Appendix 1, for the full research protocol).

We collected 391 articles from Web of Science (WoS) and Scopus databases between 2011 and 2023, mapping a 12-year span. Since we aimed to have a broad understanding of the concept and trace its evolution over time, we chose to start our analysis in 2011. This year marked the first significant mention of data literacy in education and social sciences journals, based on a preliminary review of the databases, capturing the period when data literacy began gaining prominence. We only included English, peer-reviewed, open- or early-access journal articles and book chapters in the field of social and education sciences. Review articles, proceedings, essays, and non-academic publications were not included. Due to accessibility, consistency, and resource constraints, we did not include non-English language sources, with the exception of translated sources. Our search string—("Data literac*") AND ("competence" OR "skills" OR "proficiency" OR "understanding" OR "use")—was too limiting, only giving 104 results starting in 2016. Accordingly, we broadened our search looking for papers containing "Data literac*." We placed no other restrictions on article selection. The search results were exported into Excel and duplicates were manually removed.

After applying our inclusion criteria and the removal of the duplicates (103 studies were duplicates), a selection of 288 studies' titles and abstracts was reviewed in-depth. Fifty-two studies were removed that did not mention data literacy in the title or abstract and did not address and/or lacked emphasis on data literacy competencies. Studies that fell outside of the education and social sciences were also excluded. Another 26 results were removed as they did not fit the formal article criteria (proceeding papers, reviews, etc.). After applying these criteria, 210 articles remained for analysis, as illustrated in Figure 1. The full research protocol in which all decisions are explained in detail can be found in the Supplementary File. The coding was conducted by the main author and checked by the secondary author. Cases where there was doubt were discussed amongst the author team and a decision was made.

The metadata from WoS and Scopus were exported in Excel, charting the following data: author(s), article title, abstract, publication year, source title (journal), author keywords, publisher, publisher's address, research areas, and DOI. For further analysis, several categories were added in new columns, namely "topic," "subtopic 1," "subtopic 2," "target group(s)/demographic," and "specific concepts."

As part of our mixed methods approach, we applied a quantitative frequency analysis using 4CAT Capture and Analysis Toolkit (Peeters & Hagen, 2021)—a tool for natural language processing—as well as a bibliometric analysis. The quantitative frequency analysis, after tokenising all the words, reviewed the word frequencies in titles and article abstracts allowing us to note the most common topics and chart their presence over a 12-year span. For the bibliometric analysis, we checked the main journals, authors, and their geographical distribution. This was followed by a qualitative analysis applying grounded theory with three

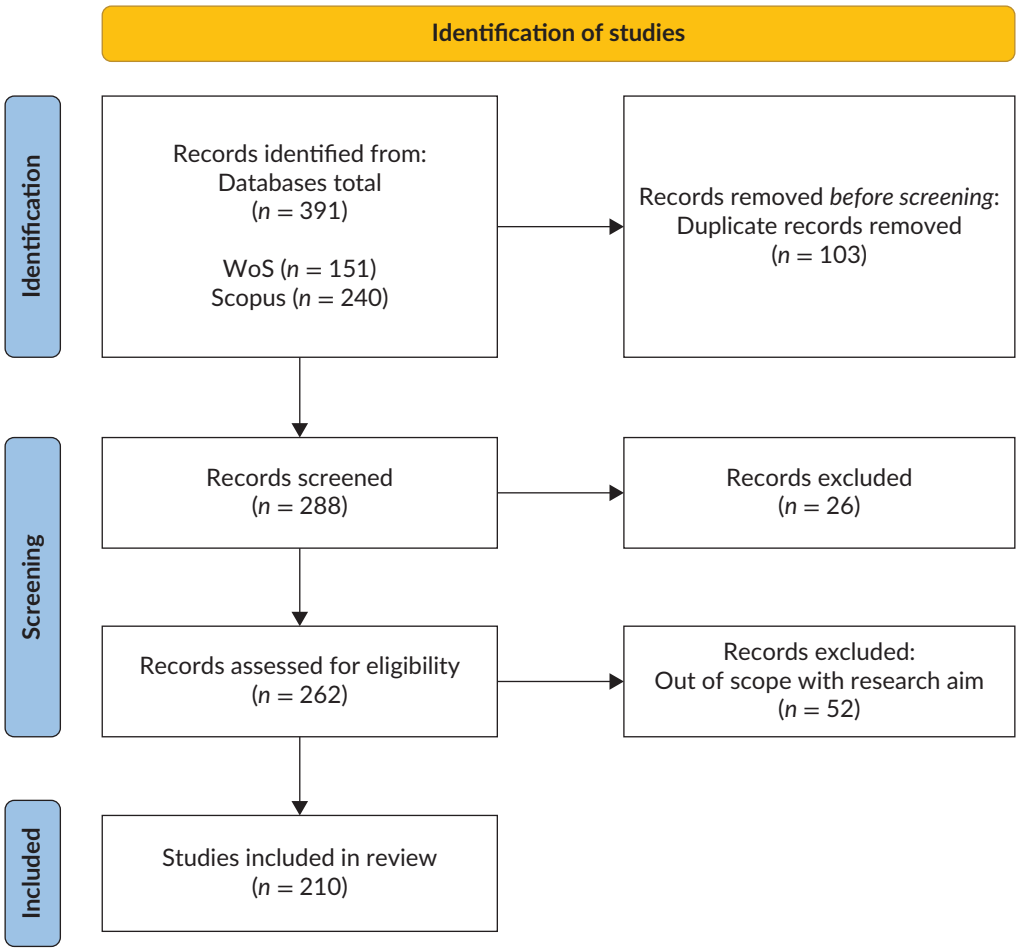

Figure 1. Prisma flowchart.

levels of thematic coding (open, axial, and selective; Glaser & Strauss, 1967; Strauss & Corbin, 1990). In this process, sensitising concepts were used, using the most frequent topics discovered through the word frequency analysis to guide our thematical coding process to allow for easier comparison at the end (Bowen, 2006). Concretely, the existing codes based on the quantitative analysis were used as inspiration during the analysis. However, room was allowed for an inductive approach to discover emerging themes and topics not captured by the quantitative analysis, leading to new codes (Bryman, 2016).

3. Discussion of Results

3.1. Bibliometric Analysis of the Literature

This section analyses the bibliometric data of the 210 selected articles. By examining the evolution of publications from 2011 to 2023, identifying the most prolific authors, and highlighting the journals with the highest frequency of publications, we aim to uncover trends in the academic discourse surrounding data literacy.

3.1.1. Evolution of the Number of Articles Published

During 2011–2023, we observe that the number of published articles on data literacy in social and education sciences (n = 210) has increased significantly (see Figure 2). The data reflects a general upward trend, with spikes in 2019, 2020 and 2023. This steady increase indicates a growing interest in and focus on data literacy within these fields over the years. Moreover, data literacy is starting to distinguish itself from other digital literacies. Despite the overall upward trend, we could attribute the dip in publications in 2021 to the possible impact of the Covid-19 pandemic, which provided challenging work circumstances for most sectors. The pandemic also influenced work performance in academia, with female academics especially experiencing a greater influence of the pandemic on their research productivity (Püttmann & Thomsen, 2024; see also Peetz et al., 2022; Sawert & Keil, 2021; Staniscuaski et al., 2021). The dip could also be attributed to academics shifting their research focus to Covid-19-related topics (Aristovnik et al., 2020; Liu et al., 2022; Roychowdhury et al., 2022), although this is hard to say definitively (Sohrabi et al., 2021). Overall, the broader trend implies a continuously growing interest in data literacy.

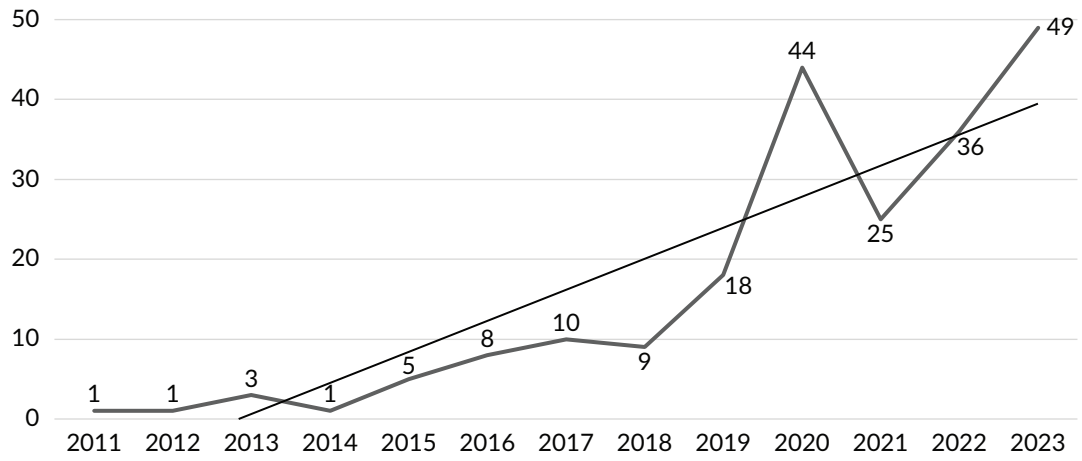

Figure 2. Evolution of number of publications per year (n = 210).

3.1.2. Most Occurring Journals

To gain an overview of the journals with the most publications (Figure 3), we conducted a bibliometric analysis of the journals (n = 135). Most articles were published in a 2020 special issue of the *Journal of Media Literacy Education*, the journal of the National Association for Media Literacy Education (n = 7, 5.18%). This was followed by the *Journal of Business and Finance Librarianship* (n = 6, 4.44%; Routledge, Taylor & Francis), of which five articles were published in 2020, and *Teaching and Teacher Education* (n = 6, 4.44%; Elsevier), which included publications ranging from 2015 to 2022. *Information and Learning Sciences* (Emerald Publishing) published five articles (3.7%). Two journals, *British Journal of Educational Technologies* (Wiley) and *Education and Information Technologies* (Springer), had four articles each (2.96%). Other journals had three or fewer publications on the topic.

Although a select number of journals had a slightly higher number of publications (Figure 3), the top-occurring journals had only a few articles published on data literacy. This indicates that the research on data literacy is spread across various disciplines and journals and is not concentrated in a single source, although we must

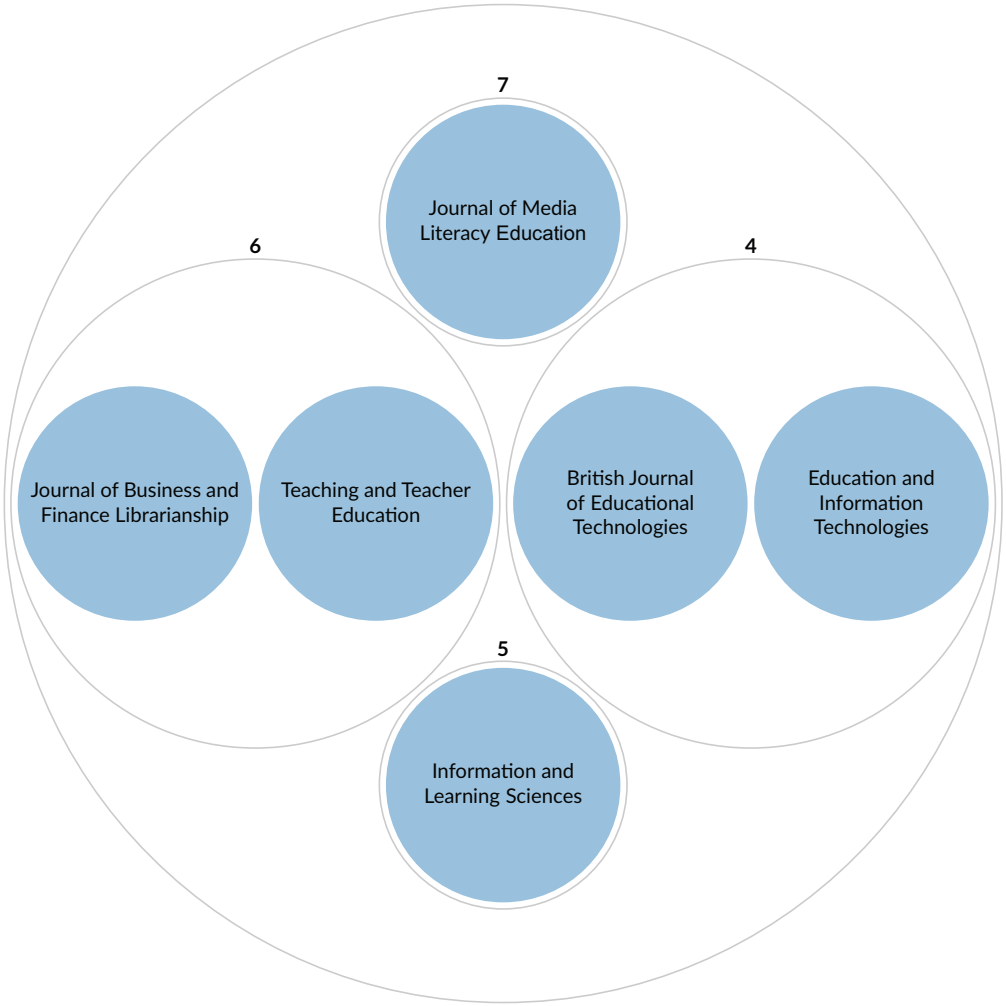

Figure 3. Top occurring journals ($n = 210$).

also consider that only open-access articles were selected for analysis. Overall, this shows that whilst research on data literacy is distributed across many journals, a few journals publish more data literacy-related content, with a notable peak in specialised publications around 2020.

3.1.3. Top Published Authors and Geographical Distribution

This subchapter provides an overview of the author diversity in the selected publications as well as the most prolific authors (Figure 4) and the geographical location of their most current research institution. We note that, in terms of author diversity, 575 authors are involved in the 210 publications. Fifty publications (23.8%) are single-authored contributions. The maximum number of authors in one publication is nine authors. Of the 210 publications, 433 authors (75.3%) are involved in only one, while 142 authors are involved in at least two. A few authors are more prolific in publishing research within the field of data literacy, with Raffaghelli J. E. ($n = 6$), Italy, and Shreiner T. L. ($n = 5$), US, standing out as the leading contributors. Beck J. S., US, and Cowie B., New Zealand, (co-)authored four papers. This was followed by Atenas J. (UK), Burress T. (US), Condon P. (US), Edwards F. (New Zealand), Havemann L. (UK), Koltay T. (Hungary), Mandinach E. B. (US), McGowan B. S. (US), Nguyen D. (The Netherlands), Reeves T. D. (US), and Whitesides H. (US) each (co-)authoring three papers.

Figure 4. Top occurring authors (n = 210).

The geographical distribution of top authors shows that they are all based in countries commonly referred to as the West, and primarily in the US.

3.2. Thematic Trends and Topical Distribution in Data Literacy Research

In this section, we aim to unpack the thematic trends within data literacy research, identifying key areas that have garnered attention and how these topics intersect. Through a word frequency analysis combined with inductive qualitative coding, we identified six main categories under which the diverse topics in data literacy research can be grouped: "teaching and student learning," "professional development and capacity building," "data use and knowledge-based decision making," "critical thinking and ethics," "research and analytical skills," and "engagement and society," although most topics are interconnected.

The word frequency analysis of the selected articles highlights the evolution of frequent topics in data literacy research (2011–2023), with Figure 5 visualising certain topical trends. It portrays an emerging and growing field, where earlier publications reflect broader discussions about data literacy. Within the field, there is a distinct interest in incorporating data literacy into education and training settings, though various

trends and understandings on how to go about this are observed. We also note a diversification of themes over time, such as an increasing frequency of "critical," "ethics," and "social," suggesting an emerging focus on a critical and ethical approach to data literacy, alongside analytical skills. The terms "professional" and "development" highlight a focus on building capacity for educators to effectively use, understand and teach data. The connections in the graph (Figure 5) suggest that data literacy is increasingly being explored beyond the central discussion around education. This shows the complex, nuanced field of data literacy with connected concepts and themes, showing the interconnections between categories. For instance, discussions on data-based decision-making frequently intersect with professional development and analytical skills. Nevertheless, the categorisation highlights a broad range of topics and a multifaceted field.

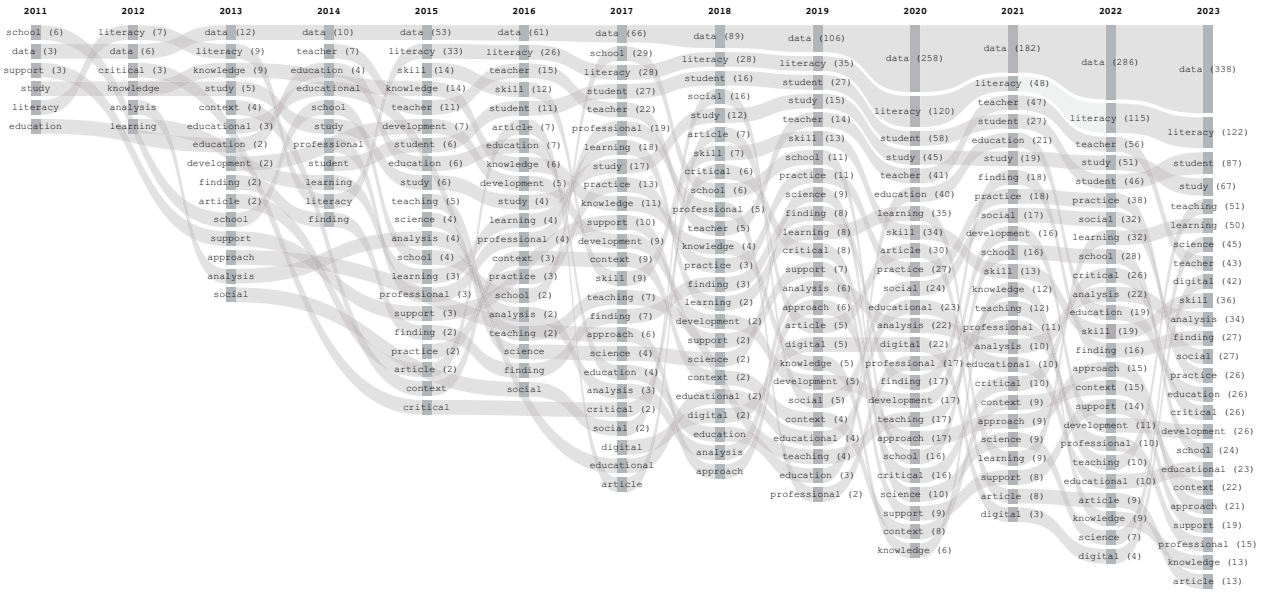

Figure 5. Evolution of 25 most occurring concepts ($n = 210$).

3.2.1. Demographic Focus in Data Literacy Publications

Regarding the addressed demographics in the selected publications, we observe that most articles discuss and research educational contexts, thus developing frameworks, interventions, and other for educators, students, and researchers. A significant portion of the content, programs, or discussions are focused on students in formal education settings. This includes various subcategories including but not limited to undergraduate students, graduate students, K-12 students, and secondary school students, with a particular emphasis on higher education. There is also a primary focus on educators and formal schooling, covering various types of teachers (preservice, novice, elementary, secondary, and tertiary education) and schools in general. This was an expected outcome due to the focus on social and education sciences. Still, the results reflect a broad focus on supporting teaching staff and improving educational settings and a significant focus on the professional development of (preservice) educators, although the focus of these articles mainly seems to remain on technical skills (Carey et al., 2018; Cowie & Cooper, 2017; Edwards et al., 2022). This is confirmed by the systematic literature review by Raffaghelli and Stewart (2020), who identified a strong focus on technical and operational skills within higher education, with the authors arguing for a critical approach to educators' data literacy development. Libraries and the role of librarians garner less focus. However, as the third most common category with a more specialised focus on enhancing the role of

libraries as educational support services, libraries are seen as essential in providing access to resources and supporting data literacy (Copeland et al., 2021; Pothier & Condon, 2020; Seidlmayer et al., 2020).

The focus on these three subgroups suggests a research gap for other target groups and sectors. Governments, civil society actors, and non-formal education lack representation in the selected literature. There is also less focus on the private sector and specific data professionals, although data literacy is crucial to work equitably and inclusively. Likewise, there is less attention on other demographics, such as older adults, parents, and vulnerable communities, although several articles define data literacy as a transversal and even lifelong skillset (Atenas et al., 2020; O'Neill, 2019; Seidlmayer et al., 2020).

3.2.2. Trends in the Selected Literature

Earlier publications concerning data literacy, such as Berendt (2012), Prado and Marzal (2013), and MacMillan (2015), focus on data literacy from an information perspective. For instance, Berendt (2012) defines data and privacy literacy as important sub-competences for information literacy. Data skills have frequently been discussed under the broader umbrella of information literacy, which includes the ability to navigate and use different types of information in a digital environment, encapsulating competences from computer, media, and data literacy (Bawden, 2001; Schield, 2004). Information literacy emphasises competences on searching for and assessing information, which are essential for understanding and using data (Berendt, 2012; Schield, 2004).

In addition, earlier publications primarily focused on the relevance of data use for development, for example using assessment data to inform instruction (Piro et al., 2014; Vanhoof et al., 2011). That said, there is a continuous interest in data-based decision-making and development, especially within the school context. Authors call for improved data literacy for teachers to handle and analyse assessment data to adapt their instruction and/or improve the school. These articles highlight the value of data-driven and evidence-based decision-making and tend to go hand in hand with calls for professional development and capacity building among educational staff (Bocala & Boudett, 2015; Conn et al., 2022; Cowie & Cooper, 2017; Hansen & Wasson, 2016; Luo et al., 2022; Piro et al., 2014). This is something Carey et al. (2018) refer to as "data-driven pedagogy." Yet, Whitesides and Beck (2020, p. 1) note that teachers are not data literate enough to "recognize inequitable data practices," allowing room for data bias in their teaching. Gibson et al. (2023, p. 1) find that "to expand students' data skills, educators must be grounded in a framework that holistically considers ethics, community impact, and science." In turn, we also find a focus on resource and support development for educators to effectively teach data literacy, such as introducing "colleague coaching" (Edwards et al., 2022).

Alongside this discussion, several authors also call for an updated education curriculum for students (Gehrke et al., 2021; Wilkerson et al., 2022), emphasising the importance of including data science and/or data-related practices, with a focus on statistical thinking, practical application, and inter-/multidisciplinary approach due to a cultural shift referred to as the "data culture" by Gould (2021). This could lead to improved student understanding and appreciation of the practical uses of statistics, preparing students to handle real-world scenarios (Bailey et al., 2020; Ow-Yeong et al., 2023; Watson & Smith, 2022). Adding to this, LaMar and Boaler (2021) reiterate the need for a curricula update to include "data science." The authors note that this empowers students by providing them with tools to analyse and understand the data they

encounter during their everyday lives whilst also stressing the importance of working with authentic, real-world data and being "critical consumers" of this data. Researchers emphasise hands-on, experiential learning approaches to build students' data literacy in practical and engaging ways (Cottone et al., 2023; Halliday, 2019; Kjelvik & Schultheis, 2020), but observe challenges in implementing data literacy education, such as curriculum constraints, disciplinary focus, and the need for resources (Ow-Yeong et al., 2023; Yoon & Copeland, 2020).

Especially after 2019, we find a growing attention for critical thinking, ethical data use, and civic participation. The increase in attention to critical thinking tends to go hand in hand with authors touching upon the importance of civic participation (or citizenship), bias in data, and transparency of data use (Gutiérrez, 2019; Heiser et al., 2023; McCosker, 2022; McCosker et al., 2022; Taylor et al., 2024; Whitesides & Beck, 2020). Markham (2019) observed critical thinking as crucial for data literacy, so that those who are data literate not only know how to use and critique data but are also able to take "social action." Sander (2020) argues for "critical big data literacy," focusing on awareness and reflection rather than data use. Whereas the understanding of the term can be inconsistent, "critical data literacy" is a concept mentioned in several abstracts, arguing that the traditional definitions of data literacy focus too much on data analytics and that the societal and participatory aspect is not considered sufficiently in models and frameworks.

The need for citizens to understand their data surroundings is not only mentioned by authors focusing on data science and statistics. There is a steady growth of articles focusing on citizen empowerment and awareness (Calabrese Barton et al., 2021; Carmi et al., 2020; Gebre, 2018; Pangrazio & Sefton-Green, 2020; Raffaghelli, 2020; Raffaghelli et al., 2024). A data-literate citizenry can create societal change, for instance by using open data for community development (Yoon & Copeland, 2020)—if the data is accessible and citizens have the necessary resources. For this, several authors call for working collaboratively with authentic, or real-world, data, for example developing climate change solutions by working with open data (Kuhn et al., 2023; see also Robertson & Tisdall, 2020; Palsa & Mertala, 2023; Werning, 2020). Open data is a growing concept within data literacy publications and often emerges as a subtheme within both education and social sciences, supporting our decision to treat it as an interdisciplinary concept rather than a distinct field. For instance, in education, open data is often integrated into discussions on data-driven learning, while in social sciences, it is connected to topics such as community engagement, accessibility, and transparency. Especially as of 2020, we observe a rise in publications that note its importance for data advocacy and societal change through data. Loría Solano et al. (2023) published a systematic literature review on open data literacy, noting a lack of open data literacy and a focus on technical skills in literature.

Carmi et al. (2020) refer to this as "data citizenship." Bhargava et al. (2022) find that data has become central to our democratic society, and to have influence means to understand data. They argue that traditional approaches to teaching data literacy, which often emphasise computational and statistical methods and skills, may not be sufficient to engage a broader and more diverse group of learners. To ensure wider participation and understanding in civic settings, it is necessary to rethink how data literacy is introduced and taught in schools, for example through data theatre and a focus on social justice (Bhargava et al., 2022), as there is a need for broader engagement. This echoes the discussion around critical data literacy. Data literacy education could benefit from embracing more interactive and creative learning techniques. Werning (2020) introduces the concept of making data "playable" by using elements of play and games to interpret real-world datasets. This approach aims to engage students in a more interactive and creative

way, contrasting with traditional and statistical methods that often focus heavily on visual evidence. By integrating other learning strategies, such as games, into data analysis, the article suggests a new approach to improving creative data literacy. This method encourages students to explore and interpret data in novel ways, potentially leading to deeper and more innovative understandings of data (Werning, 2020). This perspective complements ongoing discussions about broadening data literacy to encompass a wider range of skills and approaches and finding diverse ways to engage learners (Cottone et al., 2023; Sanei et al., 2024). Additionally, it allows for a wider range of individuals to connect with data, not only those who have a strong background in statistics and an interest in data analysis. This could foster a more inclusive learning environment that ensures all participants get to develop their data competences engagingly and meaningfully.

The last trend we wish to highlight is a continuous interest in data usage and (learning) analytics and data management (Ahmad et al., 2019; McGowan et al., 2023; Mosha & Ngulube, 2023; Robinson & Bawden, 2017; Vilar & Zabukovec, 2019). Within an information library and research context, articles on this topic tend to study how librarians and other data professionals can support students, researchers, and those in a data-rich environment in managing and analysing data (Condon & Pothier, 2022; Copeland et al., 2021; Giudice da Silva Cezar & Maçada, 2021). This is also referred to as "data librarianship" (Pothier & Condon, 2020), with libraries being considered valuable "information centres" (Seidlmayer et al., 2020) to support the community's data needs. MacMillan (2015) argues that librarians need to "expand" their skillset to be able to use and instruct about data resources, libraries being a valuable resource for individuals to gain the necessary data (research) skills. Likewise, public libraries are seen by users as important places for accessing data and for educational programs that teach data skills, especially in ways that are relevant to the local community. Therefore, libraries can adapt and build on this to meet the users' needs and tailor their data services to help users improve their data literacy (Copeland et al., 2021).

Libraries can play a significant role in helping communities access, understand, and use open data. Libraries and other information centres/organisations can serve as vital hubs for accessing, learning about, and using open data, and other types of user support. Thereby empowering individuals and communities to make informed decisions, support innovation, and engage more fully in civic life. The inclusive approach mentioned above aligns well with the role of public libraries and community centres as key spaces for accessing and learning about data. By adopting this approach, libraries can enhance their educational programs and data services, ensuring they are relevant and accessible to the local community and meet users' diverse needs.

3.2.3. Identification of Diverse Data Literacy Competences and Interpretations

Even in earlier publications on data literacy, authors referred to the importance of critical thinking (Berendt, 2012) and democratic education (Perrotta, 2013) in relation to data literacy, as well as linking it to information literacy (Perrotta, 2013; see also Berendt, 2012; Koltay, 2015; MacMillan, 2015). However, most articles primarily touched upon competences related to using, or more specifically, analysing data (Piro et al., 2014; Vanhoof et al., 2011). As of 2015, there are also mentions of science education, the alignment of data literacy and research, and the importance of finding (MacMillan, 2015) and managing one's data (Koltay, 2015, 2016). Mandinach and Gummer (2016, p. 366) lay out "the specific knowledge, skills, and dispositions teachers need to use data effectively and responsibly," calling for data literacy curricula integration. The integration of data literacy in curricula has been advocated by numerous authors over the

years. There remains a significant need across various countries to distinguish data literacy from media literacy and to integrate it into primary, secondary, higher, and adult education programs.

As of 2016, we observe the first mention of "data visualisations" in the selected literature. Philip et al. (2016) link the ability to interpret and develop data visualisations to civic participation. The article highlights a project that aimed to address the lack of data literacy in education by incorporating lessons on data collection, analysis, interpretation, representation, visualisation, and communication into a computer science class, emphasising the importance of understanding data in context. Specifically, data literacy needs to incorporate racial literacy, ensuring that students can examine race and power dynamics in data visualisations. This highlights the possible societal influence of data visualisations and narratives. Shreiner (2018, 2020) touches upon something similar, zooming in on the presence and purpose of data visualisations in social studies textbooks. We observe that in this context, data literacy involves more than just the technical skills to interpret and manage data and includes recognising how data might reinforce or challenge existing racial stereotypes and inequalities. It requires an understanding of the societal, political and cultural context, considering power dynamics, possible data risks, and the influence of data narratives (Donohoe & Costello, 2020; Kouts-Klemm, 2019; Mueller, 2022). Nonetheless, most articles in 2016–2017 remain focused on using data for development and data usage education (Reeves & Chiang, 2019; Stowell Bracke, 2017; van Geel et al., 2017). The earlier definitions of data literacy seem to primarily revolve around using data, specifically on data analytic skills and open dispositions, with a couple of outliers that influenced later conceptual discussions.

As of 2018, we see a steady increase of critical thinking and reflection-related competences addressed, as mentioned in Section 3.2.2. Moving beyond technical proficiency, several abstracts highlight how data education often focuses too narrowly on quantitative data skills, neglecting the broader understanding of data's role in everyday life (Gebre, 2018; Gray et al., 2018). For example, as Burns and Matthews (2018) argue, data literacy in journalism should extend beyond knowing how to work with data technically. It involves ethical, reflective, and critical thinking needed to authenticate, analyse, and communicate data wisely, effectively and responsibly. Authors recognise data as powerful tools for communication and advocate for a more nuanced and integrated approach to data literacy education that prepares individuals to navigate and critically engage with the data surrounding them and even influence the systems that produce and manage data (Gray et al., 2018). Data literacy is a form of influence and power. Therefore, a critical data citizenry would be able to navigate their data surroundings, hold data users accountable and advocate for their rights by demanding transparency and ethical data use (Dander & Macgilchrist, 2022; Fotopoulou, 2021; Hagen, 2022).

Carmi et al. (2020) emphasised the importance of developing data literacy initiatives that go beyond the individual, incorporate critical thinking about the online ecosystem, and empower citizens to be proactive. This perspective aligns with ongoing discussions about the need for more socially aware approaches to data literacy. As of 2020, this discussion on civic engagement and data literacy slowly but increasingly focuses on social justice and data risks and highlights power and accountability, although the focus remains on using data, though no longer merely for school improvement. Data literacy is currently seen as a potential instrument to promote social justice, although there is no unified approach (Raffaghelli, 2020). Nguyen and Beijnon (2024) highlight how critical data literacy can empower users to challenge and even resist the manipulative practices of big tech. Data literacy not only provides individuals with the tools to navigate

society but could also play a significant role in shaping activism, as well as personal and scientific inquiry. The need for transparency and the challenges posed by vague data practices echo the broader concerns about data literacy and power dynamics explored in other articles.

The gaps between the current state of data literacy education and the more comprehensive approaches advocated in the most recent literature are evident. While there is increasing awareness of the need to broaden data literacy to include socio-technical, ethical, and critical dimensions, much work remains to be done to integrate these elements into educational curricula and training. Interventions and educational projects often still emphasise data usage and less data understanding, especially when aimed at educators.

3.2.4. Shift in Focus Towards Societal Awareness and Engagement

The abovementioned topical evolution is similar to that of digital and media literacy. Early on, research on these literacies primarily focused on the proficiency to use digital technology and media systems, only later evolving to a more creative, active, and critical stance, combining competences that touch upon both using and understanding (Iordache et al., 2017; van Dijk, 2020). More recently, since the widespread use of social media, literature has been referencing "new media literacy," the definition adapting to the rise of, for instance, social media and information-seeking and sharing (Celik et al., 2021). The concept argues the importance of a critical outlook in an era of "new media technologies," providing a competence model that could allow users to recognise and deal with misinformation (Celik et al., 2021). This relates closely to the evolution of data literacy. A growing attention on "critical thinking," "awareness," and "citizenship" implies the growing importance of proactive, critical citizens who can navigate changes in the information society, such as the ever-growing influence of algorithms and AI.

Academics note that the previous definitions of data literacy, which were based on statistics and computational literacy, mainly focused on data analytics and data use for decision-making processes, leaving little room for critical analysis from sociology and community perspectives. Over the years this has evolved from a traditional data skills approach to models and definitions that also include critical thinking and data equity. Comparably, Mandinach and Gummer (2016) developed a data literacy framework for teachers, with the primary focus on data use for teaching improvement. However, in later publications, Mandinach and Jimerson (2021, p. 9) expand that data use is "at the heart of data ethics," educators need to carefully consider their actions regarding data usage, the methods they employ in their work, and how they focus their efforts on benefiting—rather than harming—the stakeholders involved. Atenas et al. (2023, p. 1) developed an "ethical framework" for data literacy in research, ensuring that their framework teaches an ethical approach to research data "to enable a critical understanding of the techno-centric environment and the intersecting hierarchies of power embedded in technology and data." These examples show an evolution from technical skills towards not only critical thinking, but also data ethics, although this trend remains relatively small, indicating a need for further exploration. The focus remains predominantly on data usage; however, considering a critical or community-focused perspective has become widespread.

4. Conclusion

In this study, we reviewed the state of data literacy research and the concept's topical evolution (2011–2023) by examining the main trends and topics. After selecting and coding 210 academic publications,

we found that the research on data literacy is ever-evolving. Similar to media literacy (Iordache et al., 2017), the concept of data literacy has grown increasingly complex, encompassing a wider range of competences over time. In response to the need for greater clarity and unity in the field, much of the research has focused on lesson interventions, educational projects, and the development and testing of models and frameworks for teaching data literacy. This goes hand in hand with the field recognising data literacy as a multifaceted, transversal skillset crucial for civic participation.

We observe that data literacy began to separate itself from information literacy and expanded beyond data science and statistics education to include the competences required to understand one's data surroundings and a critical mindset when it comes to using data (Gebre, 2018; Selwyn & Pangrazio, 2018). Research indicates that the challenge is not merely about teaching individuals to work with data but about equipping them with the skills and understanding to engage with data as informed individuals. This research field now evolves from a focus on individual data skills into data literacy for citizens capable of shaping a datafied society in ethical and socially responsible ways, thus moving from consumers to creators (Bhargava et al., 2022; Raffaghelli et al., 2024; Sander, 2024).

Similarly, most of the research published focuses on teaching technical data skills and learning analytics for development, albeit increasingly critical (Heiser et al., 2023; Mandinach & Jimerson, 2021). Attention for equitable data use and data misuse challenges also garners more attention in these areas, however, it needs to be emphasised that awareness is crucial for preventing the misuse of data and for fostering informed, responsible data use. This is increasingly studied within the school context and is even starting to be explored within the public sector (Clegg et al., 2023; Fotopoulou, 2021; Nguyen, 2023).

Although we observe an increase in critical thinking within data literacy publications, there is still room for research that touches upon the ethical use of data, specifically in the subfield of data justice and data advocacy. Particularly in community- and citizen-focused research, research could address the broader societal impacts of data-driven decisions. Libraries, as non-formal learning environments, are seen as valuable resources for skill development and data access as well as community engagement (Copeland et al., 2021; Seidlmayer et al., 2020). The influence of other information centres and civil society organisations on community development and citizen empowerment leaves room to be studied and possibly compared to or complement the role of libraries.

We also find a lack of attention to digital accessibility and inclusion in the research, even though there is an increase in agency-related topics, such as data justice, empowerment through data literacy, and data for equity. Therefore, further research is needed on data literacy in vulnerable communities and other understudied demographics. As argued in multiple publications, all individuals need to possess the necessary skills and knowledge to navigate a datafied society effectively. Accordingly, data literacy programs and resources should be designed with inclusivity and social support in mind. It is crucial that these programs not only focus on developing practical data usage but also address the potential dangers of data misuse and data risks (Atenas et al., 2023; Dodman et al., 2023; Mandinach & Jimerson, 2021), particularly within the private sector, which remains understudied.

This study also has its limitations. The scoping review identified key themes and knowledge gaps for future research; however, it mainly provides a surface-level overview based on the article titles and abstracts.

Due to this article providing a summary of the most notable trends in data literacy research, it remains quite descriptive, providing a synthesis of themes and patterns rather than certain nuances or complexities around data literacy. The variability of studies included in the research also made data analysis challenging. Scoping reviews, unlike systematic literature reviews, do not tend to include a quality check of the articles included nor an assessment of intervention effectiveness. To address these limitations, we only included peer-reviewed publications (2011–2023). Additionally, this review only covers open- and early-access English-language articles indexed in the WoS and Scopus databases, leaving room for future research to include other databases and national citation indexes. Our decision to focus on English-language articles was made for practical reasons, but we acknowledge the inherent language bias this introduces. To mitigate this limitation, we sourced articles from two international databases and employed both automated searches and manual screening. While these measures reduce the risk of algorithmic bias, we recognise that some degree of bias remains unavoidable. In addition, the scoping review provides a basis for more in-depth systematic literature reviews.

Research into specific concepts such as AI, algorithms, and privacy within the broader framework of data literacy is a topic for further research, as there was a lack of focus on it in the selected abstracts. Due to the growing interest in these topics, focus on social and education research might shift towards these topics' role in data literacy. This could also help to explore critical issues such as transparency and accountability in AI and algorithms. This is essential for understanding how data literacy intersects with various technological advancements and ethical concerns, further emphasising its societal relevance.

More research is required to foster a data literacy that is meaningful and applicable in real-world scenarios and engaging for all citizens, not merely those with an aptitude for statistics. This correlates with the need to have more effect and impact measurement research on the interventions and resources developed. Data literacy represents a dynamic paradigm, continually evolving as societal, technological, and educational contexts shift. Therefore, it needs to be tailored to address the varying levels of engagement and expertise of different demographics, as they encounter different needs and challenges. Similar to digital literacy, data literacy does not require a one-size-fits-all approach (Mensonides et al., 2024), but a nuanced and context-specific approach, providing a socially aware and adaptable data literacy that includes social support structures. This adaptability reflects the complex nature of the field, with new challenges and opportunities reshaping the approach to data literacy.

Acknowledgments

We would like to express our heartfelt gratitude to our colleague Ibe Delvaux for her invaluable guidance in setting up the scoping review process, and to our colleague Anna Luo for her meticulous proofreading of our article. The LLM ChatGPT was used to assist in the writing process of this article. Specifically, the tool was asked several times for grammar and sentence structure improvement.

Funding

This article is part of the PhD scholarship titled From Classrooms to Communities? Data Literacy in Formal and Non-Formal Education, conducted at imec-SMIT, Vrije Universiteit Brussel.

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Ahmad, K., JianMing, Z., & Rafi, M. (2019). Librarian's perspective for the implementation of big data analytics in libraries on the bases of lean-startup model. *Digital Library Perspectives*, *36*(1), 21–37. https://doi.org/10.1108/DLP-04-2019-0016

Ahmed, Y., Walker, C., & Leon-Urrutia, M. (2021). Defining data literacy communities by their objectives: A text mining analysis. In O. Senevirtane, V. Singh, A. Freire, J.-D. Luo (Eds.), *WebSci '21 Companion: Companion Publication of the 13th ACM Web Science Conference 2021* (pp. 23–26). Association for Computing Machinery. https://doi.org/10.1145/3462741.3466663

Aristovnik, A., Ravšelj, D., & Umek, L. (2020). A bibliometric analysis of COVID-19 across science and social science research landscape. *Sustainability*, *12*(21), Article 9132. https://doi.org/10.3390/su12219132

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Atenas, J., Havemann, L., & Timmermann, C. (2020). Critical literacies for a datafied society: Academic development and curriculum design in higher education. *Research in Learning Technology*, *28*, Article 2468. https://doi.org/10.25304/rlt.v28.2468

Atenas, J., Havemann, L., & Timmermann, C. (2023). Reframing data ethics in research methods education: A pathway to critical data literacy. *International Journal of Educational Technology in Higher Education*, *20*(1), Article 11. https://doi.org/10.1186/s41239-023-00380-y

Bailey, J., Cowie, B., & Cooper, B. (2020). "Maths outside of maths": Pre-service teachers' awareness of mathematical and statistical thinking across teachers' professional work. *Australian Journal of Teacher Education*, *45*(1), 1–18. https://doi.org/10.14221/ajte.2020v45n1.1

Bawden, D. (2001). Information and digital literacies: A review of concepts. *Journal of Documentation*, *57*(2), 218–259. https://doi.org/10.1108/EUM0000000007083

Berendt, B. (2012). Data mining for information literacy. In D. E. Holmes & L. C. Jain (Eds.), *Data mining: Foundations and intelligent paradigms* (pp. 265–297). Springer. https://doi.org/10.1007/978-3-642-23151-3_12

Bhargava, R., Brea, A., Palacin, V., Perovich, L., & Hinson, J. (2022). Data theatre as an entry point to data literacy. *Educational Technology & Society*, *25*(4), 93–108. https://www.jstor.org/stable/48695984

Bhargava, R., Deahl, E., Letouzé, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). *Beyond data literacy: Reinventing community engagement and empowerment in the age of data*. Data-Pop Alliance & Internews.

Bocala, C., & Boudett, K. P. (2015). Teaching educators habits of mind for using data wisely. *Teachers College Record: The Voice of Scholarship in Education*, *117*(4), 1–20. https://doi.org/10.1177/016146811511700409

Bowen, G. A. (2006). Grounded theory and sensitizing concepts. *International Journal of Qualitative Methods*, *5*(3), 12–23. https://doi.org/10.1177/160940690600500304

Bryman, A. (2016). *Social research methods* (5th ed.). Oxford University Press.

Burns, L. S., & Matthews, B. J. (2018). First things first: Teaching data journalism as a core skill. *Asia Pacific Media Educator*, *28*(1), 91–105. https://doi.org/10.1177/1326365X18765530

Calabrese Barton, A., Greenberg, D., Turner, C., Riter, D., Perez, M., Tasker, T., Jones, D., Herrenkohl, L. R., & Davis, E. A. (2021). Youth critical data practices in the COVID-19 multipandemic. *AERA Open*, *7*. https://doi.org/10.1177/23328584211041631

Carey, M., Grainger, P., & Christie, M. (2018). Preparing preservice teachers to be data literate: A Queensland case study. *Asia-Pacific Journal of Teacher Education*, *46*(3), 267–278. https://doi.org/10.1080/1359866X.2017.1402860

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. https://dx.doi.org/10.1353/pla.2011.0022

Carmi, E., Yates, S. J., Lockley, E., & Pawluczuk, A. (2020). Data citizenship: Rethinking data literacy in the age of disinformation, misinformation, and malinformation. *Internet Policy Review*, *9*(2). https://doi.org/10.14763/2020.2.1481

Celik, I., Muukkonen, H., & Dogan, S. (2021). A model for understanding new media literacy: Epistemological beliefs and social media use. *Library & Information Science Research*, *43*(4), Article 101125. https://doi.org/10.1016/j.lisr.2021.101125

Clegg, T. L., Cleveland, K., Weight, E., Greene, D., & Elmqvist, N. (2023). Data everyday as community-driven science: Athletes' critical data literacy practices in collegiate sports contexts. *Journal of Research in Science Teaching*, *60*(8), 1786–1816. https://doi.org/10.1002/tea.21842

Condon, P. B., & Pothier, W. G. (2022). Advancing data literacy: Mapping business data literacy competencies to the ACRL framework. *Journal of Business & Finance Librarianship*, *27*(2), 104–126. https://doi.org/10.1080/08963568.2022.2048168

Conn, C. A., Bohan, K. J., Bies-Hernandez, N. J., Powell, P. J., Sweeny, S. P., Persinger, L. L., & Persinger, J. D. (2022). Expected data literacy knowledge and skills for early career teachers: Perspectives from school and district personnel. *Teaching and Teacher Education*, *111*, Article 103607. https://doi.org/10.1016/j.tate.2021.103607

Copeland, A., Yoon, A., & Zhang, S. (2021). Data reuse practices and expectations for data resources and services among public library users. *Public Library Quarterly*, *40*(4), 330–345. https://doi.org/10.1080/01616846.2020.1773749

Cottone, A. M., Yoon, S. A., Shim, J., Coulter, B., & Carman, S. (2023). Evaluating the apt epistemic processes of data literacy in elementary school students. *Instructional Science*, *51*(1), 1–37. https://doi.org/10.1007/s11251-022-09610-8

Cowie, B., & Cooper, B. (2017). Exploring the challenge of developing student teacher data literacy. *Assessment in Education: Principles, Policy & Practice*, *24*(2), 147–163. https://doi.org/10.1080/0969594X.2016.1225668

Dander, V., & Macgilchrist, F. (2022). School of data and shifting forms of political subjectivity. In P. Bettinger (Ed.), *Educational perspectives on mediality and subjectivation* (pp. 45–67). Springer. https://doi.org/10.1007/978-3-030-84343-4_3

D'Ignazio, C. (2017). Creative data literacy: Bridging the gap between the data-haves and data-have nots. *Information Design Journal*, *23*(1), 6–18. https://doi.org/10.1075/idj.23.1.03dig

Dodman, S. L., DeMulder, E. K., View, J. L., & Stribling, S. M. (2023). "I knew it was a problem before, but did I really?": Engaging teachers in data use for equity. *Journal of Educational Change*, *24*, 995–1023. https://doi.org/10.1007/s10833-022-09477-z

Donohoe, D., & Costello, E. (2020). Data visualisation literacy in higher education: An exploratory study of understanding of a learning dashboard tool. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(17), Article 115. https://doi.org/10.3991/ijet.v15i17.15041

Edwards, F., Cowie, B., & Trask, S. (2022). Using colleague coaching to develop teacher data literacy. *Professional Development in Education*. Advance onlne publication. https://doi.org/10.1080/19415257.2022.2081247

Fotopoulou, A. (2021). Conceptualising critical data literacies for civil society organisations: Agency, care, and social responsibility. *Information, Communication & Society*, *24*(11), 1640–1657. https://doi.org/10.1080/1369118X.2020.1716041

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. https://doi.org/10.1007/s42330-018-0034-z

Gehrke, M., Kistler, T., Lübke, K., Markgraf, N., Krol, B., & Sauer, S. (2021). Statistics education from a data-centric perspective. *Teaching Statistics*, *43*(1), 201–215. https://doi.org/10.1111/test.12264

Ghodoosi, B., Torrisi-Steele, G., West, T., & Li, Q. (2023). An exploration of the definition of data literacy in the academic and public domains. *International Journal of Adult Education and Technology*, *14*(1), 1–16. https://doi.org/10.4018/IJAET.325218

Gibson, H., Short, M. E., & O'Donnell, C. (2023). Data, discourse, and development: Building a sustainable world through education and science communication. *Frontiers in Communication*, *8*, Article 981988. https://doi.org/10.3389/fcomm.2023.981988

Giudice da Silva Cezar, B., & Maçada, A. C. G. (2021). Data literacy and the cognitive challenges of a data-rich business environment: An analysis of perceived data overload, technostress and their relationship to individual performance. *Aslib Journal of Information Management*, *73*(5), 618–638. https://doi.org/10.1108/AJIM-01-2021-0015

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Sociology Press.

Gould, R. (2021). Toward data-scientific thinking. *Teaching Statistics*, *43*(1), 11–22. https://doi.org/10.1111/test.12267

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, *5*(2). https://doi.org/10.1177/2053951718786316

Gutiérrez, M. (2019). Participação num ambiente datificado: Questões sobre literacia de dados. *Comunicação e Sociedade*, *36*, 37–55. https://doi.org/10.17231/comsoc.36(2019).2342

Hagen, A. N. (2022). Datafication, literacy, and democratization in the music industry. *Popular Music and Society*, *45*(2), 184–201. https://doi.org/10.1080/03007766.2021.1989558

Halliday, S. D. (2019). Data literacy in economic development. *The Journal of Economic Education*, *50*(3), 284–298. https://doi.org/10.1080/00220485.2019.1618762

Hansen, C. J., & Wasson, B. (2016). Teacher inquiry into student learning:—The TISL heart model and method for use in teachers' professional development. *Nordic Journal of Digital Literacy*, *11*(1), 24–49. https://doi.org/10.18261/issn.1891-943x-2016-01-02

Heiser, R., Stritto, M. E. D., Brown, A., & Croft, B. (2023). Amplifying student and administrator perspectives on equity and bias in learning analytics: Alone together in higher education. *Journal of Learning Analytics*, *10*(1), 8–23. https://doi.org/10.18608/jla.2023.7775

Iordache, E., Mariën, I., & Baelden, D. (2017). Developing digital skills and competences: A QuickScan analysis of 13 digital literacy models. *Italian Journal of Sociology of Education*, *9*(1), 6–30.

Kjelvik, M. K., & Schultheis, E. H. (2020). Getting messy with authentic data: Exploring the potential of using data from scientific research to support student data literacy. *CBE—Life Sciences Education*, *18*(2), 1–8. https://doi.org/10.1187/cbe.18-02-0023

Koltay, T. (2015). Data literacy: In search of a name and identity. *Journal of Documentation*, *71*(2), 401–415.

Koltay, T. (2016). Data governance, data literacy and the management of data quality. *IFLA Journal*, *42*(4), 303–312. https://doi.org/10.1177/0340035216672238

Kouts-Klemm, R. (2019). Data literacy among journalists: A skills-assessment based approach. *Central European Journal of Communication*, *12*(3), 299–315. https://doi.org/10.19195/1899-5101.12.3(24).2

Kuhn, C., Pete, J., & Raffaghelli, J. E. (2023). Critical data literacies for good. In L. Czerniewicz & C. Cronin (Eds.), *Higher education for good* (1st ed., pp. 491–508). Open Book Publishers. https://doi.org/10.11647/obp.0363.21

LaMar, T., & Boaler, J. (2021). The importance and emergence of K-12 data science. *Phi Delta Kappan*, *103*(1), 49–53. https://doi.org/10.1177/00317217211043627

Liu, Y.-L., Yuan, W.-J., & Zhu, S.-H. (2022). The state of social science research on COVID-19. *Scientometrics*, *127*(1), 369–383. https://doi.org/10.1007/s11192-021-04206-4

Louie, J., Stiles, J., Fagan, E., Chance, B., & Roy, S. (2022). Building toward critical data literacy with investigations of income inequality. *Educational Technology & Society*, *25*(4), 142–163.

Luo, J., Wang, M., & Yu, S. (2022). Exploring the factors influencing teachers' instructional data use with electronic data systems. *Computers & Education*, *191*, Article 104631. https://doi.org/10.1016/j.compedu.2022.104631

MacMillan, D. (2015). Developing data literacy competencies to enhance faculty collaborations. *Liber Quarterly*, *24*(3), 140–160.

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. https://doi.org/10.1016/j.tate.2016.07.011

Mandinach, E. B., & Jimerson, J. B. (2021). Data ethics in education: A theoretical, practical, and policy issue. *Studia Paedagogica*, *26*(4), 9–26. https://doi.org/10.5817/SP2021-4-1

Markham, A. N. (2019). Critical pedagogy as a response to datafication. *Qualitative Inquiry*, *25*(8), 754–760. https://doi.org/10.1177/1077800418809470

McCosker, A. (2022). Making sense of deepfakes: Socializing AI and building data literacy on GitHub and YouTube. *New Media & Society*, *26*(5), 2786–2803. https://doi.org/10.1177/14614448221093943

McCosker, A., Yao, X., Albury, K., Maddox, A., Farmer, J., & Stoyanovich, J. (2022). Developing data capability with non-profit organisations using participatory methods. *Big Data & Society*, *9*(1). https://doi.org/10.1177/20539517221099882

McGowan, B. S., Ekeigwe, A., & Clase, K. (2023). Designing and assessing a data literacy internship program for graduate health sciences students. *Journal of the Medical Library Association*, *110*(4), 501–506. https://doi.org/10.5195/jmla.2022.1498

Mensonides, D., Smit, A., Talsma, I., Swart, J., & Broersma, M. (2024). Digital literacies as socially situated pedagogical processes: Genealogically understanding media, information, and digital literacies. *Media and Communication*, *12*, Article 8174. https://doi.org/10.17645/mac.8174

Mosha, N. F., & Ngulube, P. (2023). Teaching research data management courses in higher learning institutions in Tanzania. *Library Management*, *44*(1/2), 166–179. https://doi.org/10.1108/LM-04-2022-0033

Mueller, J. (2022). Data visualization and advocacy for sexual and gender minority health. *Journal of Map & Geography Libraries*, *18*(1/2), 115–130. https://doi.org/10.1080/15420353.2022.2139329

Nguyen, D. (2023). How news media frame data risks in their coverage of big data and AI. *Internet Policy Review*, *12*(2). https://doi.org/10.14763/2023.2.1708

Nguyen, D., & Beijnon, B. (2024). The data subject and the myth of the 'black box' data communication and critical data literacy as a resistant practice to platform exploitation. *Information, Communication & Society*, *27*(2), 333–349. https://doi.org/10.1080/1369118X.2023.2205504

O'Neill, S. A. (2019). Fomento de la alfabetización de datos en todos los ámbitos. *Aula Abierta*, *48*(4), 419–434. https://doi.org/10.17811/rifie.48.4.2019.419-434

Ow-Yeong, Y. K., Yeter, I. H., & Ali, F. (2023). Learning data science in elementary school mathematics: A comparative curriculum analysis. *International Journal of STEM Education*, *10*(1), Article 8. https://doi.org/10.1186/s40594-023-00397-9

Palsa, L., & Mertala, P. (2023). Contextualizing everyday data literacies: The case of recreational runners. *International Journal of Human–Computer Interaction*, *40*(19), 5845–5856. https://doi.org/10.1080/10447318.2023.2241612

Pangrazio, L., & Sefton-Green, J. (2020). The social utility of 'data literacy.' *Learning, Media and Technology*, *45*(2), 208–220. https://doi.org/10.1080/17439884.2020.1707223

Peeters, S., & Hagen, S. (2021). The 4CAT Capture and Analysis Toolkit: A modular tool for transparent and traceable social media research. *Computational Communication Research*, *4*(2), 571–589. https://doi.org/10.2139/ssrn.3914892

Peetz, D., Baird, M., Banerjee, R., Bartkiw, T., Campbell, S., Charlesworth, S., Coles, A., Cooper, R., Foster, J., Galea, N., de la Harpe, B., Leighton, C., Lynch, B., Pike, K., Pyman, A., Ramia, I., Ressia, S., Samani, M. N., Southey, K., . . . Weststar, J. (2022). Sustained knowledge work and thinking time amongst academics: Gender and working from home during the COVID-19 pandemic. *Labour and Industry*, *32*(1), 72–92. https://doi.org/10.1080/10301763.2022.2034092

Perrotta, C. (2013). Assessment, technology and democratic education in the age of data. *Learning, Media and Technology*, *38*(1), 116–122. https://doi.org/10.1080/17439884.2013.752384

Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction*, *34*(4), 361–388. https://doi.org/10.1080/07370008.2016.1210418

Piro, J. S., Dunlap, K., & Shutt, T. (2014). A collaborative data chat: Teaching summative assessment data use in pre-service teacher education. *Cogent Education*, *1*(1), Article 968409. https://doi.org/10.1080/2331186X.2014.968409

Pothier, W. G., & Condon, P. B. (2020). Towards data literacy competencies: Business students, workforce needs, and the role of the librarian. *Journal of Business & Finance Librarianship*, *25*(3/4), 123–146. https://doi.org/10.1080/08963568.2019.1680189

Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, *63*(2), 123–134.

Püttmann, V., & Thomsen, S. L. (2024). Academics' susceptibility to disruptions of their research productivity: Empirical insights from the COVID-19 pandemic. *Higher Education*, 1–21. https://doi.org/10.1007/s10734-024-01266-2

Raffaghelli, J. E. (2020). Is data literacy a catalyst of social justice? A response from nine data literacy initiatives in higher education. *Education Sciences*, *10*(9), Article 233. https://doi.org/10.3390/educsci10090233

Raffaghelli, J. E., Romero Carbonell, M., & Romeu-Fontanillas, T. (2024). In my opinion, the TOS….Situating personal data literacy interventions. *Information and Learning Sciences*, *125*(5/6), 367–388. https://doi.org/10.1108/ILS-06-2023-0086

Raffaghelli, J. E., & Stewart, B. (2020). Centering complexity in 'educators' data literacy' to support future practices in faculty development: A systematic review of the literature. *Teaching in Higher Education*, *25*(4), 435–455. https://doi.org/10.1080/13562517.2019.1696301

Reeves, T. D., & Chiang, J.-L. (2019). Effects of an asynchronous online data literacy intervention on pre-service and in-service educators' beliefs, self-efficacy, and practices. *Computers & Education*, *136*, 13–33. https://doi.org/10.1016/j.compedu.2019.03.004

Robertson, J., & Tisdall, E. K. M. (2020). The importance of consulting children and young people about data literacy. *Journal of Media Literacy Education*, *12*(3), 58–74. https://doi.org/10.23860/JMLE-2020-12-3-6

Robinson, L., & Bawden, D. (2017). "The story of data": A socio-technical approach to education for the data librarian role in the CityLIS library school at City, University of London. *Library Management*, *38*(6/7), 312–322. https://doi.org/10.1108/LM-01-2017-0009

Roychowdhury, K., Bhanja, R., & Biswas, S. (2022). Mapping the research landscape of Covid-19 from social sciences perspective: A bibliometric analysis. *Scientometrics*, *127*(8), 4547–4568. https://doi.org/10.1007/s11192-022-04447-x

Salomão Filho, A., Wasson, B., Lameras, P., & Maloszek, R. (2023). Alfabetización de datos para la ciudadanía: Formulación de algunas recomendaciones políticas a través de una revisión sistemática. *Edutec. Electronic Journal of Educational Technology*, (86), 6–21. https://doi.org/10.21556/edutec.2023.86.2877

Sander, I. (2020). What is critical big data literacy and how can it be implemented? *Internet Policy Review*, *9*(2). https://doi.org/10.14763/2020.2.1479

Sander, I. (2024). Critical datafication literacy—A framework for educating about datafication. *Information and Learning Sciences*, *125*(3/4), 270–292. https://doi.org/10.1108/ILS-06-2023-0064

Sanei, H., Kahn, J. B., Yalcinkaya, R., Jiang, S., & Wang, C. (2024). Examining how students code with socioscientific data to tell stories about climate change. *Journal of Science Education and Technology*, *33*(2), 161–177. https://doi.org/10.1007/s10956-023-10054-z

Sawert, T., & Keil, M. (2021). Research note: The COVID-19 pandemic and its effects on scientific work in German sociology. *Zeitschrift Für Soziologie*, *50*(5), 338–348. https://doi.org/10.1515/zfsoz-2021-0023

Schield, M. (2004). Information literacy, statistical literacy and data literacy. *IASSIST Quarterly*, *28*(8), 6–11. https://doi.org/10.29173/iq790

Seidlmayer, E., Müller, R., & Förstner, K. U. (2020). Data literacy for libraries—A local perspective on library carpentry. *Bibliothek Forschung Und Praxis*, *44*(3), 485–489. https://doi.org/10.1515/bfp-2020-2038

Selwyn, N., & Pangrazio, L. (2018). Doing data differently? Developing personal data tactics and strategies amongst young mobile media users. *Big Data & Society*, *5*(1). https://doi.org/10.1177/2053951718765021

Seymoens, T., Van Audenhove, L., Van den Broeck, W., & Mariën, I. (2020). Data literacy on the road: Setting up a large-scale data literacy initiative in the DataBuzz project. *Journal of Media Literacy Education*, *12*(3), 102–119. https://doi.org/10.23860/JMLE-2020-12-3-9

Shreiner, T. L. (2018). Data literacy for social studies: Examining the role of data visualizations in K–12 textbooks. *Theory & Research in Social Education*, *46*(2), 194–231. https://doi.org/10.1080/00933104.2017.1400483

Shreiner, T. L. (2020). Data-literate citizenry: How US state standards address data and data visualizations in social studies. *Information and Learning Sciences*, *121*(11/12), 909–931. https://doi.org/10.1108/ILS-03-2020-0054

Sohrabi, C., Mathew, G., Franchi, T., Kerwan, A., Griffin, M., Del Mundo, J. S. C., Ali, S. A., Agha, M., & Agha, R. (2021). Impact of the coronavirus (COVID-19) pandemic on scientific research and implications for clinical academic training—A review. *International Journal of Surgery*, *86*, 57–63. https://doi.org/10.1016/j.ijsu.2020.12.008

Solano, E. L., Catasús, M. G., & Raffaghelli, J. E. (2023). (Open) data literacy: Which relationships with open data adoption? A systematic review of the literature. *Italian Journal of Educational Technology*, *31*(2), 37–55. https://doi.org/10.17471/2499-4324/1303

Staniscuaski, F., Kmetzsch, L., Soletti, R. C., Reichert, F., Zandonà, E., Ludwig, Z. M. C., Lima, E. F., Neumann, A., Schwartz, I. V. D., Mello-Carpes, P. B., Tamajusuku, A. S. K., Werneck, F. P., Ricachenevsky, F. K., Infanger, C., Seixas, A., Staats, C. C., & de Oliveira, L. (2021). Gender, race and parenthood impact academic productivity

during the COVID-19 pandemic: From survey to action. *Frontiers in Psychology*, *12*, Article 663252. https://doi.org/10.3389/fpsyg.2021.663252

Stowell Bracke, M. (2017). Agricultural librarians becoming informationists: A paradigm shift. *Library Trends*, *65*(3), 448–466. https://doi.org/10.1353/lib.2017.0011

Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Sage.

Taylor, M., Kennedy, H., & Oman, S. (2024). Challenging assumptions about the relationship between awareness of and attitudes to data uses amongst the UK public. *The Information Society*, *40*(1), 32–53. https://doi.org/10.1080/01972243.2023.2283729

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Van Audenhove, L., Van den Broeck, W., & Mariën, I. (2020). Data literacy and education: Introduction and the challenges for our field. *Journal of Media Literacy Education*, *12*(3), 1–5. https://doi.org/10.23860/JMLE-2020-12-3-1

Van Audenhove, L., Vermeire, L., Van den Broeck, W., & Demeulenaere, A. (2024). Data literacy in the new EU DigComp 2.2 framework how DigComp defines competences on artificial intelligence, internet of things and data. *Information and Learning Sciences*, *125*(5/6), 406–436. https://doi.org/10.1108/ILS-06-2023-0072

van Dijk, J. (2020). *The digital divide*. Polity.

van Geel, M., Keuning, T., Visscher, A., & Fox, J.-P. (2017). Changes in educators' data literacy during a data-based decision making intervention. *Teaching and Teacher Education*, *64*, 187–198. https://doi.org/10.1016/j.tate.2017.02.015

Vanhoof, J., Verhaeghe, G., Verhaeghe, J. P., Valcke, M., & Van Petegem, P. (2011). The influence of competences and support on school performance feedback use. *Educational Studies*, *37*(2), 141–154. https://doi.org/10.1080/03055698.2010.482771

Vilar, P., & Zabukovec, V. (2019). Research data management and research data literacy in Slovenian science. *Journal of Documentation*, *75*(1), 24–43. https://doi.org/10.1108/JD-03-2018-0042

Watson, J., & Smith, C. (2022). Statistics education at a time of global disruption and crises: A growing challenge for the curriculum, classroom and beyond. *Curriculum Perspectives*, *42*(2), 171–179. https://doi.org/10.1007/s41297-022-00167-7

Werning, S. (2020). Making data playable: A game co-creation method to promote creative data literacy. *Journal of Media Literacy Education*, *12*(3), 88–101. https://doi.org/10.23860/JMLE-2020-12-3-8

Whitesides, H., & Beck, J. S. (2020). "There is subjectivity, there is bias": Teacher candidates' perceptions of equity in data literacy for teaching. *The Teacher Educator*, *55*(3), 283–299. https://doi.org/10.1080/08878730.2020.1725792

Wilkerson, M. H., Lanouette, K., & Shareff, R. L. (2022). Exploring variability during data preparation: A way to connect data, chance, and context when working with complex public datasets. *Mathematical Thinking and Learning*, *24*(4), 312–330. https://doi.org/10.1080/10986065.2021.1922838

Williamson, B., Bayne, S., & Shay, S. (2020). The datafication of teaching in higher education: Critical issues and perspectives. *Teaching in Higher Education*, *25*(4), 351–365. https://doi.org/10.1080/13562517.2020.1748811

Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, *12*(3), 9–26. https://doi.org/10.15353/joci.v12i3.3275

Wuyckens, G., Landry, N., & Fastrez, P. (2022). Untangling media literacy, information literacy, and digital literacy: A systematic meta-review of core concepts in media education. *Journal of Media Literacy Education*, *14*(1), 168–182. https://doi.org/10.23860/JMLE-2022-14-1-12

Yoon, A., & Copeland, A. (2020). Toward community-inclusive data ecosystems: Challenges and opportunities of open data for community-based organizations. *Journal of the Association for Information Science and Technology*, *71*(12), 1439–1454. https://doi.org/10.1002/asi.24346

About the Authors

Lotte Vermeire is a joint PhD researcher in communication sciences at imec-SMIT, Vrije Universiteit Brussel (VUB) in Belgium and information systems at the University of the Western Cape (UWC) in South Africa. She is part of the Digital Inclusion and Citizen Engagement and Media, Marketing, and User Experience units. Her research looks into data literacy, digital youth work, and digital inclusion. Her PhD project focuses on digitally inclusive initiatives aimed at improving data literacy in both formal and non-formal educational settings.

Wendy Van den Broeck (PhD) is a senior lecturer at the Department of Communication Sciences (VUB) and heads the Media, Marketing, and User Experience Unit (MUX). She is also program director of the bachelor Social Sciences and teaches several courses within the Communication and Social Sciences programs at VUB. Her expertise relates to user research on new media, personalised media, and digital practices. She co-founded DataBuzz, for which she received the Royal Flemish Academy of Belgium for Science and the Arts' (KVAB) annual prize for science communication.

Fazlyn Petersen (PhD) is a senior lecturer in the Department of Information Systems at the University of the Western Cape (UWC). Her research foci are information communication and technology for development (ICT4D) in health and education. Her research focuses on creating more inclusive online environments for students and patients, especially those with lower socioeconomic status.

Leo Van Audenhove (PhD) is full professor at the Department of Communication Sciences (VUB) and is extra-ordinary professor at UWC. In 2013, he took the lead in setting up the Knowledge Centre for Digital and Media Literacy in Flanders, of which he subsequently became the director. The centre was established by the government as an independent centre to promote digital and media literacy in Flanders. His research focuses on media literacy, digital inclusion and ICT4D. In 2021–2022 Leo held a Francqui chair at the Université de Namur.