Associated conference: “Yes we can!” - Digital Education for Better Futures (EDEN 2023 Annual Conference)

Conference location: Dublin City University (DCU), Dublin, Ireland

Conference date: 18-20 June 2023


Published on: 27 October 2023

Copyright: © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.
EMPOWERING DIVERSE LEARNERS THROUGH DESIGN-BASED DATA LITERACY EDUCATION: THE DATA DESIGN CYCLE FRAMEWORK

Nico Kling, University of Applied Sciences Zwickau, Germany
Chantal Kling, University of Applied Sciences Zwickau, Germany
Emelie Schwill, University of Applied Sciences Zwickau, Germany
Christian-Andreas Schumann, University of Applied Sciences Zwickau, Germany
Sajal Kabiraj, Häme University of Applied Sciences, Finland

Correspondence: Emelie Schwill: emelie.schwill.gfb@fh-zwickau.de

Abstract
Data literacy is a vital life-skill that is becoming increasingly necessary in today’s society. As such, public data literacy education must become a substantive component in teaching, requiring an appropriate pedagogical model. This paper aims to present a new definition for data literacy, as well as conceptual framework for public data literacy education, drawing from tenets of design-based education and synthesized principles of data literacy education from literature. The resulting framework is termed the Data Design Cycle, and emphasizes collaboration, communication, and the development of problem-solving skills. The model is highly adaptable to different educational environments and learning styles, making it easily integrable into a variety of educational settings, including traditional classroom-based learning, online learning, and experiential learning programs. Furthermore, by promoting the principles of public teaching of data literacy and a learner-centred approach, the Data Design Cycle ensures that data literacy education is accessible to heterogeneous learners from diverse educational and socioeconomical settings.

Keywords:
Data Literacy; Design-based Education; Future Competence; Life Skill; Diversity in Education

Data Literacy is a life skill
According to Einay and Levin (2014), data is now more accessible and has broader applications, including new observations and dimensions. However, despite the opportunities presented by big data analytics, there is widespread public skepticism that it threatens privacy, restricts civil liberties, and leads to increased state and corporate control (Koltay, 2015). This skepticism is heightened with the increasing complexity of data, which has led to a diminished role of human factors (Wolff et al., 2016). Therefore, the importance of understanding data has become more critical than ever, particularly in the ethical use of data (Carlson et al., 2011). Datafication, the transformation of social actions into quantifiable data, has become a new paradigm in society and science (Van Dijck, 2014), leading to data becoming a valuable commodity and emphasizing the importance of mastering data skills as a life skill. The term commonly used to describe these data skills is data literacy (Koltay, 2015). The concept of data literacy is defined in various ways, with definitions typically categorized into general and discipline-dependent, and further divided into technical and sociotechnical approaches. Technical definitions focus on the technical processes behind data literacy, emphasizing the academic skills required to handle information enriched with computer science and statistics (Raffaghelli, 2020). In contrast, sociotechnical definitions expand on technological processes and include cognitive processes such as data ethics and critical thinking. One perspective describes it as the existence of data awareness and sensitivity in order to effectively and appropriately collect, analyse, process, use and present data (Bai et al., 2021). We combined these categorizations into a new definition:
Data Literacy is a purpose-bound collection of various sociological and technical competencies, consisting of knowledge, skills and behaviour, that build on each other. The sum of the competencies allows the holistic handling of data, which in its most pronounced form encompasses the entire data lifecycle from the creation to the disposal.

Our reason for defining data literacy as a purpose-bound collection of sociological and technical competencies, rather than a list of specific knowledge, abilities, and behaviours, is that we believe this distinction is necessary for the construct to be universally valid. A purpose-bound definition recognizes that data literacy is not a fixed set of skills or knowledge, but rather a set of competencies that are developed and used in different contexts to achieve specific goals. By emphasizing the purpose-bound nature of data literacy, we hope to create a more comprehensive and flexible definition that can be applied to diverse educational settings and real-world contexts.

We believe that in a datafied society, public access to data literacy education is substantive. Only then can the citizens use their knowledge to contribute to the positive development of society. However, the current state of data literacy education in mainly done in formal educational settings, and is often inadequate, as evidenced by the low levels of proficiency demonstrated by students and professionals alike (e.g. Ogugu et al., 2020; Zhao, 2020). This highlights the need for a review of existing literature and pedagogical approaches to data literacy in education, as it allows us to identify principles that are applicable in diverse educational settings, including public education, and better equip learners with the necessary data literacy skills for the future. Additionally, due to data literacy not only being a technical skill but also a mindset and a way of thinking, traditional teaching methods may not be enough to fully equip learners with the necessary competencies, especially in the public sphere. Thus, we additionally introduce a hybrid approach that combines the problem-solving and user-centred design aspects of design-based education based on Geitz & de Geus (2019) with the technical aspects of data literacy education which could be the key to effectively teaching data literacy skills publicly. Our conceptual framework is termed Data Design Cycle.

Therefore, this paper seeks to address the question: “How can the principles of design-based education and data literacy be applied to create a comprehensive and adaptable framework for teaching data literacy in diverse educational settings?”

Five Key Principles for Fostering Data Literacy through Education

When it comes to developing a comprehensive understanding of a complex topic such as data literacy in education, an explorative approach to a literature review can be highly valuable. By reviewing a broad range of existing data literacy education research and synthesizing the key findings into five broadly applicable principles for effective data literacy education, we can develop a more nuanced and holistic understanding of the topic.

The first principle we found can be derived from the various definitions of data literacy in the existing literature, which is primarily classified as a competence, skill, or ability within an educational context. This classification is significant, as a competence is a more comprehensive term than the other two. As Marzal (2020) notes, a competence is linked to an attitude and a behaviour that allows for successful completion of an activity and efficient execution of activities, with transferability to various areas of application. However, many of the existing definitions do not fully consider the mindset aspect of data literacy (e.g. Burress et al., 2020; Bautista et al., 2017; Raffaghello, 2020; Fraser-Amott, 2020). As per the definition, a competent individual must be able to apply the behaviors in various interdisciplinary settings, which is also missing from many definitions. The understanding of data literacy as a competence provides a foundation for a holistic framework that includes all aspects of the term, and is crucial in determining the choice of teaching content and curriculum creation. This leads to Principle 1:

1. Principle: Understanding the concept of competence is important for teaching data literacy.

The second principle that we derived from our explorative approach is the importance of teachers being proficient in data literacy themselves in order to effectively teach the subject. The works of Starobin & Upah (2014) and Yang & Li (2020) highlight that data exists in universities and can be utilized for teaching through learning analytics and other technologies. However, the authors argue that to effectively apply these teaching-enhancing technologies, teachers must have a thorough understanding of data literacy. By becoming competent in data literacy, teachers can integrate data-driven insights into their pedagogical practices and develop innovative teaching strategies that better meet the needs of their students. Therefore, we propose the following principle:
2. Principle: Teachers need to possess a strong foundation in data literacy to deliver effective data literacy education.

The literature furthermore shows that learners are more likely to have a greater understanding of data literacy when they are actively involved in the learning process. Gläser & Spree (2022) and Werning (2020) suggest that students are more engaged and motivated when the teaching aligns with their reality. Maybee & Zilinski (2015) proposed a data-based teaching approach that can be used to achieve this alignment. The approach involves using study objects that are representative of the educational reality in the students' chosen discipline. The same idea applies to the choice of the underlying definition for the teaching intervention, which should be aligned with the learners' discipline depending on the teaching objectives (Atenas et al., 2020; Hitchcock et al., 2021; Starobin & Upah, 2014; Todorova et al., 2019; Werning, 2020). Therefore, the third principle that can be derived from these arguments is:

3. Principle: Effective data literacy education should be based on the learners' reality and aligned with their discipline.

Drawing from the literature, it is clear that the practical skills required for data literacy are just as important as knowledge and education culture. Examples of successful practical teaching methods are discussed in the literature (Schüller, 2022; Gläser & Spree, 2022; Werning, 2020; Gehrke et al., 2021; Hitchcock et al., 2021; Cheng et al., 2019), and can be virtual or analog in nature. Building practical competence not only empowers students but also encourages their involvement in the public sphere. This can be achieved through public lectures, workshops, and other initiatives that foster cooperation among learners, teachers, and society. Such engagement strengthens democracy, promotes social justice, and provides economically applicable competences (Raffaghelli, 2020; Atenas et al., 2020). In light of these findings, the fourth principle for effective data literacy education can be articulated as follows:

4. Principle: Data literacy education should be conducted in a practical and publicly engaged manner.

To synthesize effective data literacy education, it is essential to recognize that data literacy is not a skill that can be learned in isolation. Rather, it is a complex and long-term process that most likely requires a comprehensive curriculum to be taught effectively, specifically in higher education. The work of Bai et al. (2021) and Burress et al. (2020) highlights the importance of developing a structured curriculum for teaching data literacy generically. Stand-alone courses with too much content are not recommended as they do not provide students with adequate time to build the necessary competencies. Instead, entire modules that stretch the learning process and provide realistic goals are a more effective way to teach data literacy (Gläser & Spree, 2022). By understanding that data literacy is a long-term process, educators can take a more strategic and intentional approach to education, which in turn will promote better learning outcomes. This leads to the fifth and final principle:

5. Principle: The teaching and learning of data literacy should be approached as a long-term process with a comprehensive and structured curriculum.

The five principles we have developed provide a suggested foundation for effective data literacy education, both applicable for formal and informal settings. However, to implement these principles in a practical and engaging way, a suitable education approach is needed. One such approach is design-based education, which emphasizes the importance of iterative problem-solving and user-centred design. By applying the principles to a design-based education framework, we can create a pedagogical approach that has the potential to be both effective and engaging for learners. Therefore, the following chapter will explore the key tenets of design-based education and demonstrate how this approach can be used to teach data literacy in a way that promotes creativity, innovation, and problem-solving.

**Design Based Education and Data Literacy**

To effectively teach learners the necessary skills to work with data, a pedagogical approach that is multifaceted is needed. It is important to note that our understanding of DBE is based on the work by Gerry Geitz and Jan de Geus (2019). The DBE approach consists of six phases that can be related to the phases of self-regulated learning. Before starting a research/learning process to approach a problem, goal setting, planning, and
managing one's motivational beliefs are necessary processes influencing the effectiveness of the execution of this research/learning process. During the research/learning process, the application of strategies to effectively approach and execute the research (i.e., learning) supports the intended outcome of the process. Finally, after the initial research steps resulting in a solution/prototype, self-reflection is needed to determine whether the intended outcome/performance is recognized and valued by knowledgeable others (i.e., working field and lecturers). Self-reflection leads back to the forethought phase that precedes the next learning effort. The diagram illustrating the six phases of the DBE approach can be seen in figure 1.

By applying the principles of understanding the concept of competence, teaching based on the reality of the learners, practical and public teaching of data literacy, and the long-term nature of data literacy education to a DBE framework, we can create a conceptual pedagogical approach that is both effective and engaging for learners, which we named the Data Design Cycle. The results are of theoretical nature, and the application of the three different self-learning phases are still to be conceptualized.

**The Data Design Cycle**

The need for an effective pedagogical approach to data literacy has led to the development of the Data Design Cycle, an adaptation of the design-based education process that incorporates the principles of effective data literacy education. The Data Design Cycle is a cyclical process that allows learners to approach data-related problems from a problem-solving perspective, emphasizing collaboration, communication, and the development of problem-solving skills. By applying the principles of design-based to the Data Design Cycle, learners can develop their own solutions to complex data-related problems in a flexible and adaptable way. Ultimately, the goal of the Data Design Cycle is to equip learners with the necessary skills and mindset to succeed in a rapidly evolving field and to promote the use of data for positive change in society. As it is based on Geitz & de Geus (2019) DBE model, it consists of six different phases, as seen in figure 2:
1. Data Problem Framing

During the Data Problem Framing, learners identify and frame a data-related problem. This includes understanding the context, scope, and complexity of the problem. As well as considering potential stakeholders and their needs. By applying Principle 1, which emphasizes the importance of understanding the concept of competence, learners gain a clear understanding of the problem and the competencies needed to solve it. The learners should also consider the ethical and social implications of their project and ensure that their proposed solutions are grounded in responsible data use.

2. User Needs Analysis

In the User Needs Analysis, learners analyse the needs and preferences of potential users of the data-related solution they want to create. This includes considering factors such as accessibility, user-friendliness, and accuracy. By applying Principle 3, which states that teaching should be based on the reality of the learners, learners gain a deeper understanding of the practical applications of data literacy and are better equipped to design solutions that meet real-world needs. This phase helps to ensure that the final data solution will meet the needs of the intended audience.

3. Data Solution Ideation

During the Data Solution phase, learners generate ideas for potential solutions to the data-related problem they identified in the previous phase. This involves brainstorming, mind mapping, and other ideation techniques. By applying Principle 4, which advocates for practical and public teaching of data literacy, learners work to develop solutions that can be practically applied in real-world settings, while publicly discussing the social and ethical implications of their proposed solutions and ensure that their ideas are grounded in responsible data use, accepted by the community.

4. Data Solution Prototyping

During the data solution prototyping, learners create a prototype of their proposed data-related solution. This could involve creating data visualizations, developing software applications, or designing user interfaces. By applying Principle 2, which emphasizes the importance of teachers being data literate, learners develop their own data literacy skills through hands-on work.

5. Data Solution Testing
In this phase, learners test their prototype with potential users and gather feedback on its effectiveness. This feedback can be used to refine the solution and improve its usability. By applying Principle 5, which recognizes the importance of the long-term nature of data literacy education, learners are to continually refine their skills and develop a deeper understanding of data-related problem-solving.

6. Data Solution Reflection

Finally, learners reflect on their design-based data literacy experience, considering what they learned, how they could improve their approach, and how they could apply their new data literacy skills in future settings. By applying the idea of critical data literacy, learners reflect on the social and ethical implications of their work and consider how data can be used to promote positive change. This ensures that learners not only develop technical data literacy skills but also a critical perspective on data use and its implications for society.

The six phases of the Data Design Cycle provide learners with a structured approach to data literacy education, emphasizing collaboration, communication, and problem-solving skills. The framework is specifically designed to be accessible and adaptable for learners at all levels of data literacy. One of the key advantages of the Data Design Cycle is its flexibility in accommodating learners of varying levels of experience with data literacy. For beginners, the Data Design Cycle provides a structured approach to learning data literacy, with clear steps and guidance to help learners gain a foundational understanding of the key concepts and principles of data literacy. The focus on problem-solving and collaboration provides an engaging and interactive learning experience, allowing beginners to learn through hands-on experience. No-code applications, as discussed by Kling et al. (2022), could be suitable and approachable alternatives to coded solutions, and would allow beginners to test the wide range of technological possibilities for data-specific practical applications without the need for prior knowledge. For intermediate learners, the Data Design Cycle offers a platform to deepen their understanding of data literacy, building upon the foundational knowledge gained in the beginner stage. The iterative nature of the Data Design Cycle allows intermediate learners to engage with complex data-related problems and develop innovative solutions. The emphasis on collaboration and communication encourages learners to work with others, building their teamwork and leadership skills. This stage is ideal for learners who are interested in developing a deeper understanding of data literacy and its applications in real-world settings, and a switch from no-code to low-code allows to hone their data skills even more. For expert learners, the Data Design Cycle provides a space to explore cutting-edge data-related problems and develop innovative solutions to complex challenges. The flexibility of the Data Design Cycle allows expert learners to engage with data-related problems at a high level, while the emphasis on collaboration and communication provides a platform for them to share their knowledge and expertise with others. This stage is ideal for learners who have a deep understanding of data literacy and are interested in pushing the boundaries of the field, as well as users who are able to code.

Furthermore, the Data Design Cycle is highly adaptable to different educational backgrounds and learning styles. The framework can be easily integrated into a variety of public educational settings, including traditional classroom-based learning, online learning, and experiential learning programs. This adaptability ensures that the framework can be used in diverse educational contexts and can be tailored to meet the specific needs of learners in different settings. By promoting the principles of public teaching of data literacy and a learner-centred approach, the framework ensures that data literacy education is accessible to heterogeneous learners from diverse educational and socioeconomical backgrounds. Publicly teaching data literacy with the Data Design Cycle can take many forms, including workshops and community outreach programs. These programs can be tailored to meet the needs of specific groups, such as low-income communities, underrepresented groups, or marginalized populations. The focus on problem-solving and collaboration allows learners to engage with data literacy in a practical and interactive way, which increases both data literacy as well as the community spirit.

Discussion

Data literacy is becoming increasingly important in today's data-driven world, and its significance is only set to grow in the coming years. The ability to understand and work with data has become essential not just for individuals and organizations, but also for governments and society as a whole. As we become more reliant on data in various aspects of our lives, from healthcare to finance to education, being able to effectively analyse and interpret data is critical for making informed decisions and driving positive change. However, data literacy must
be understood as purpose-bound. The five principles we have identified are a crucial foundation for effective data literacy education. They provide a learner-centred approach that emphasizes the importance of understanding the learners’ needs, the practical applications of data literacy, and the social and ethical implications of data use. By prioritizing these principles in data literacy education, we can ensure that learners are not just developing technical skills, but also gaining a critical perspective on data use and its role in society. Our conceptual Data Design Cycle builds on these principles by providing a structured and adaptable framework for data literacy education that can be used by learners at all levels, from beginners to experts. By incorporating the principles of design-based education the framework emphasizes problem-solving, collaboration, and the development of practical, real-world solutions. The cycle also promotes the importance of reflecting on the learning process and specifically data ethics, making it a holistic and integrated approach to data literacy education. Furthermore, the Data Design Cycle’s emphasis on publicly teaching data literacy ensures that the framework is accessible to learners from all backgrounds, promoting greater inclusivity and diversity in data literacy education with the goal of achieving a democratic data literacy education. However, it is important to note that the field of data literacy is constantly evolving, with new technologies and data practices emerging. As such, ongoing evaluation and improvement of the Data Design Cycle approach will be crucial to ensure that it remains effective and relevant for learners at all levels. An empirical study utilizing the framework is planned to adequately incorporate the three Self-Regulation Phases, and to validate the entire model.

To conclude and as advocates for education as a public good, we both believe and hope that our pedagogical data literacy education model can empower individuals of all backgrounds to engage critically and effectively with data, enabling them to actively participate in shaping a more equitable and just society.

References


Todorova, T., Krasteva, R., & Tsvetkova, E., 2019, Data Literacy and Research Data Management: The Case at ULSIT. Communications in Computer and Information Science, 989, pp. 535–544. https://doi.org/10.1007/978-3-030-13472-3_50


