Assessing improvements in participants' understanding of data through a DALI-based workshop

Rini Oktavia[1,2]*, Syarifah Meurah Yuni[1], and Ernita Dewi Meutia[3]
[1] Department of Mathematics, Universitas Syiah Kuala, Banda Aceh, INDONESIA
[2] STEM Research Center, USK, Banda Aceh, INDONESIA
[3] Department of Electrical Engineering, USK, Banda Aceh, INDONESIA
e-mail: *rini_oktavia@usk.ac.id (corresponding author)
ORCID ID: R. Oktavia: 0000-0002-5864-0858
S. M. Yuni: 0009-0003-7518-5521
E. D. Meutia: 0000-0001-6689-5116

ABSTRACT

As data increasingly influences decision-making in educational and professional settings, data literacy has become an essential competency. This study examines the effectiveness of a workshop conducted according to the Data Literacy (DALI) for Citizenship Framework, which comprises three main components: understanding data, acting on data, and engaging with data. This study focused solely on the data understanding component. The aim was to assess whether a DALI-based educational intervention could improve participants' data literacy and learning experiences. Using a pre- and post-test design, data were collected from 35 participants in Aceh, Indonesia, who completed the workshop with a 10-item questionnaire to assess understanding of data types and a follow-up semi-structured interview to explore participants' learning experiences. A paired t-test analysis revealed a significant increase in participants' understanding of data types following the workshop; however, Cohen's d effect size indicated a moderate effect. The follow-up interview showed that the workshop increased participants' awareness of the importance of data literacy for decision-making. The workshop also enhanced participants' understanding that, every day, they are acting on and engaging with data. Participants reported positive perceptions of their learning experience, suggesting that the workshop facilitated engagement, understanding, and confidence in working with data. These findings illustrate the significance of systematic, theory-based methodologies in data literacy education. The study provides empirical validation of the DALI framework as an evaluation and instructional instrument, offering evidence for its use in educational and professional development contexts and advancing research on effective data literacy pedagogy.

Keywords: Data literacy; DALI data literacy for citizenship framework; Community library.

INTRODUCTION

The traditional and straightforward definition of literacy is the ability to read and write. However, the definition of literacy is expanding alongside advances in digital information technology and the acquisition of new skills needed to comprehend, interpret, generate, and communicate in a rapidly evolving, information-rich world (UNESCO, 2019a). Depending on the perspective and context, terms such as data literacy, media literacy, digital literacy, and statistical literacy are increasingly used. Although definitions vary across fields, there is

consensus that data literacy includes the ability to collect, manage, evaluate, and use data critically (Gummer & Mandinach, 2015; Ridsdale et al., 2015; Wolff et al., 2016; Koltay, 2017; Ghodoosi et al., 2023). Most research and educational initiatives on data literacy focus on formal settings, particularly higher education and K-12 institutions (Matuk et al., 2021; Lee et al., 2024). Bowler and Shaw (2024) found a significant lack of organised informal learning activities, especially for groups such as older adults and young children. They also found that public libraries and community-based programmes present promising yet underutilised opportunities for informal data literacy development.

The current state of informal data literacy education is characterised by increasing recognition of its importance alongside formal education; however, it remains underdeveloped and inconsistently implemented (Bowler & Shaw, 2024). The scope and importance of informal data literacy education relate to the need for individuals to understand, analyse, and use data critically in everyday situations beyond traditional educational settings. This includes digital information from social media, large data systems, and various real-world data sources. Improving data literacy is essential as society becomes increasingly data-centric (Bowler & Shaw, 2024; Castañeda et al., 2024).

Informal education often adopts humanistic, socio-technical, and critical perspectives on data literacy (Castañeda et al., 2024) and uses educational methods such as gamification, arts integration, and interactive data activities to engage learners creatively (Matuk et al., 2021; Bowler & Shaw, 2024; Castañeda et al., 2024). A key challenge is that most studies using instructional interventions lack assessment methods for measuring data literacy, limiting their ability to evaluate effects on learners' competencies (Yeom, 2025; Nwagwu, 2025).

Informal data literacy education needs more structured frameworks, inclusive teaching strategies, and greater participation beyond formal education to equip individuals with essential data skills for the 21st century (Bowler & Shaw, 2024; Castañeda et al., 2024). As communities become increasingly data-driven – affecting everything from market prices to local health indicators – Community Learning Centres (CLCs) play a crucial role as informal education providers. In low-resource settings, these centres are vital for promoting lifelong learning and community empowerment (The World Bank, 2023). However, to fulfil this role, CLC facilitators must support learners in developing data literacy skills that are meaningful, ethical, and actionable (UNESCO, 2019b). Among various educational approaches, the Data Literacy for Citizenship (DALI) framework is notable for integrating technical, critical, and civic data use (Calzada Prado & Marzal, 2013; Castañeda et al., 2024). This framework uses a competency model that divides literacy into three core domains – Understanding Data, Acting on Data, and Engaging through Data – with Ethics and Privacy as a cross-cutting concern (Castañeda et al., 2024). By moving beyond technical skills to include provenance, critical interpretation, and civic participation, the DALI model aligns with current recommendations to treat data literacy as a fundamental civic skill for adults (Mandinach & Gummer, 2016; UNESCO IITE, 2021).

Most empirical studies on data literacy have been conducted in the Global North, highlighting the need for global perspectives (Bowler & Shaw, 2024). Conducting empirical studies on data literacy in the eastern region is essential to contribute to a global perspective. This study seeks to achieve that objective. In our initial effort to conduct empirical research on data literacy in the eastern region, we employed the DALI Framework, which has been validated as effective in Europe (Castañeda et al., 2024). This study aims to

determine the efficacy of the DALI Framework in the eastern region, particularly in Aceh, Indonesia.

There is an urgent need for a robust literacy movement in Indonesia. Only 25% of Indonesian students aged 15 reached level 2 or higher in reading (OECD, 2024). Level 2 pupils should be able to locate information using explicit criteria, even when these criteria are somewhat complex. They must also identify the main ideas in relatively lengthy texts and articulate the format and purpose of writing when explicitly instructed (OECD, 2024).

Forum Taman Bacaan Masyarakat (Community Reading Centre Forum; TBM) is a leading organisation in Indonesia's literacy movement. The TBM is a designated space that provides reading materials to the community, primarily to foster a passion for reading and promote lifelong learning. The TBM may be managed by either the community or the government, serving as an educational centre and information resource. Community-managed TBMs, also known as community libraries, can be categorised as Community Learning Centres (CLCs) (Novrita et al., 2025). The non-governmental organisation Forum TBM serves as a platform for assembly and coordination for the founders and managers of TBMs, literacy advocates, communities, and society to collectively advance the literacy movement in Indonesia (Forum TBM, 2024). Forum TBM envisions itself as a hub for enhancing community literacy and fostering a capable and prosperous learning society (Forum TBM, 2024). Numerous TBMs in Aceh Province are facing sustainability challenges and require support (Novrita et al., 2025). Taking all this into account, this paper aims to elucidate the impacts of a community service and research initiative dedicated to improving data literacy among the facilitators of community libraries associated with Forum TBM Aceh, using the gamification approach proposed by the DALI framework and incorporating assessment methods to evaluate data literacy.

LITERATURE REVIEW

Literacy, in its classic and simplified sense, is defined as the ability to read and write (Stokes, 2008). As societal and workplace demands have increased, the meaning of literacy has broadened beyond this foundational definition. In 2019, UNESCO stated that in a constantly changing, information-saturated world, literacy includes not only reading and writing but also the ability to identify, understand, interpret, create, communicate, and compute using printed and written materials (UNESCO, 2019). More recently, UNESCO has extended this view to encompass a wider range of capabilities, including digital literacy, media literacy, education for sustainable development, global citizenship, and job-specific competencies. Literacy skills, therefore, continue to expand and evolve as individuals increasingly engage with information and learning through digital technologies (UNESCO, 2025).

Within this broader evolution of literacy, data literacy has become increasingly important. As the importance of data grows, the ability to understand and work with it is rapidly becoming essential in today's world (Rubin & Gould, 2023). A lack of data literacy skills not only restricts an individual's ability to secure employment in an expanding knowledge-based economy but also limits their capacity to interpret and apply critical information for health, political, and financial decision-making (Cui et al., 2023). Consequently, statistical and data literacy are now recognised as essential competencies in the 21st century (Schreiter et al., 2024). Reflecting this shift, the need to develop data literacy has been widely documented in the literature (e.g., Fontichiaro & Oehrli, 2016; Ongena, 2023; Pinto et al., 2023; Pothier & Condon, 2025).

Despite its growing importance, data literacy remains context-dependent and is neither defined nor operationalised uniformly. Ghodoosi et al. (2023) examine data literacy through various educational perspectives and note that a universally accepted definition remains elusive because expectations shift across contexts. Their study highlights a discrepancy between how data literacy is framed in academic settings and the practical expectations of public sector organisations. This misalignment reflects Koltay's (2017) argument, which distinguishes general data literacy education from the specialised skills required for the professional practice of information experts. Using Ridsdale et al.'s (2015) framework to examine these differences, Ghodoosi et al. (2023) identified a clear divergence in how data is prioritised: academic competencies often emphasise the technical lifecycle of data – such as collecting, managing, and ethically analysing data – whereas businesses and community members tend to focus on transforming data into actionable knowledge for decision-making.

The contextual nature of data literacy necessitates examining how data literacy requirements vary across disciplines and application settings. In response to this need, the present project aims to enhance data literacy skills among TBM facilitators in Aceh, with the longer-term goal of transferring these skills to the community members served by the TBMs. To support a structured, citizenship-oriented approach that aligns with non-formal learning contexts, this study adopts the descriptions of data literacy summarised by the Dali: Data Literacy for Citizenship project (Castañeda et al., 2024).

The DALI project is a collaborative initiative involving five institutions: the University of Bergen, Coventry University, Friedrich-Alexander-Universitat, the University of the Balearic Islands, and the University of Murcia. It spans four countries: Germany, Norway, Spain, and the UK. Its primary objective is to enhance data literacy among citizens, starting with adults who have significant influence over the management of citizen data. The framework comprises three main components: (i) understanding data, (ii) acting on data, and (iii) engaging through data, along with a fourth overarching component, (iv) ethics and privacy, which is integrated into each of the three elements. The framework also outlines markers of advancing proficiency for each of the three main components. In this study, the focus is on the first component, Understanding Data, which encompasses three aspects: Knowledge, Awareness, and Critical Thinking (DALI Project, 2023a).

The DALI Project aims to deliver customised data literacy learning experiences for adults in non-formal educational settings. It combines networked learning with playful learning by co-creating educational games that go beyond entertainment through intentional pedagogical design. The game-based learning approach highlights key elements such as viewing failure as a learning opportunity, fostering creativity through immersive gameplay, and engaging participants in voluntary, intrinsically motivating activities. The development process follows the Transdisciplinary Model for Developing Game-Based Interventions, which balances entertainment with educational objectives (DALI Project, 2023a).

The game used in this study, Data Iceberg, is designed to teach players about different types of data and how to identify them in everyday situations. It presents scenarios that participants may have experienced and prompts them to reflect on these situations to recognise the types of data generated – either by themselves or through their interactions with related devices. The game comprises two main sections. In the first, a memory-style guessing game, players find card pairs and recall the locations of previously revealed cards. In the second, they determine what type of information each card represents and place it on their board (DALI Project, 2023d).

Importantly, the DALI games' instructional approach is adaptable across different contexts. Facilitators can adjust gameplay to suit their participants and practical needs. Before beginning Data Iceberg, facilitators decide how participants will be grouped, for example, by shared or diverse skills, by ensuring each group includes at least one participant with advanced skills, or by allowing participants to self-organise or collaborate with unfamiliar peers. During play, the facilitator supports participants by preventing fixation on a single "correct" solution. The facilitator may introduce an answer sheet, and for some cards, more than one answer may be acceptable; if a player or group provides sufficient valid arguments, their response may be considered correct. The session may then conclude with guided participant self-reflection (DALI Project, 2023c).

This study's methodological positioning is informed by broader discussions about research skill development. Mixed-method research requires researchers to cultivate a diverse skill set spanning both quantitative and qualitative approaches. Molina-Azorin (2016) noted that researchers often find it difficult to move beyond familiar methods in which they have developed comfort and expertise. In this study, the researchers' stronger quantitative background influences the methodological approach adopted. In line with Molina-Azorin (2016), expanding and refining methodological skills can strengthen the rigour of conceptual thinking. For this reason, the study employs a quasi-mixed-method research design.

A quasi-mixed-methods research design involves collecting two distinct types of data (qualitative and quantitative) and analysing them separately. There is minimal or no integration of findings or inferences across the two strands, primarily because they do not address a shared research question that would necessitate a fully mixed-methods approach. A study does not qualify as mixed-methods research when it lacks an integrated mixed-methods research question. In addition, quantitative and qualitative analyses are not combined, and inferences from each strand are not synthesised to address the research question(s) (Tashakkori et al., 2020).

METHODS

The research employed a quasi-mixed-methods design (Tashakkori et al., 2020), which minimally integrates a one-group pre-test and post-test quasi-experimental approach with a qualitative case study component. This design allows for the quantitative assessment of changes in students' competencies before and after the intervention, while simultaneously enabling a qualitative exploration of the learning process and experience. The study was guided by the following research questions:

i.   What are participants' initial data literacy levels, specifically their understanding of data, before attending the workshop?
ii. To what extent does the DALI-based workshop improve participants' data literacy, particularly their understanding of data as measured by the post-test results?
iii. How do participants perceive their learning experience during the DALI-based workshop, particularly regarding content relevance, instructional clarity, confidence development, and applicability to real-world tasks?

The study was conducted according to the following sequence:

Pre-test
The main purpose of the pre-test is to assess participants' understanding of data types before the focus group discussion (FGD) and workshop, thereby establishing a baseline. The participants are TBM Facilitators attending the FGD and workshop. The instrument used for the pre-test was a 10-item questionnaire assessing participants' understanding of data types encountered in everyday life, such as those found in social media, e-commerce, transportation, and health apps. The pre-test was administered at the venue of the FGD, the Senate Meeting Room of the Faculty of Mathematics and Natural Sciences at Syiah Kuala University. Using a Google form, an example of the translated items is shown in Figure 1. The instrument's items are provided in Appendix 1. The responses were recorded and analysed using IBM SPSS Statistics.

> To prepare for my next holiday, I use my search engine to find good driving routes in the area. My search engine saves the websites I visited, as well as clicks and mouse movements on those websites.
>
> What type of data is stored by that search engine?
> a. Visible analytical data
> b. Visible raw data
> c. Invisible analytical data
> d. Invisible raw data
> e. Functional data

Figure 1: Example of pre- and post-test items

Focus Group Discussion (FGD)
The main purpose of the FGD was to gather participants' current knowledge, attitudes, and challenges regarding data use. The participants were TBM Facilitators affiliated with Forum TBM Aceh. The instruments used were a semi-structured question guide, a Moderator Guide, and a Qualitative Coding Framework (see Appendix 2). In a group discussion moderated by a research team representative, participants were encouraged to answer open-ended questions about their experiences with data, the types of data they work with, and the challenges to data literacy. A note-taker recorded responses and observations from the discussion. The Qualitative Coding Framework was used to analyse the concepts and themes that emerged from the FGD. At the end of the FGD, the moderator summarised the key points and explained the next step: a workshop to enhance participants' data literacy.

Workshop
The main purpose of the workshop was to help participants understand data concepts and data types, enabling them to recognise each type and understand how data influences their actions in everyday life. The participants were the 35 TBM Facilitators who also attended the FGD. The instrument used was a memory board game (see a capture of the memory board game instruments in Appendix 3) adapted from the "Data Iceberg" activities developed by the DALI: Data Literacy for Citizenship Project (Castañeda et al., 2024). A mini-lecture at the beginning of the workshop explained the types of data, how each can be found in daily activities, and how they influence our lives. After the mini-lecture, a hands-on activity using the memory board game was conducted with all participants, who were divided into eight groups. The memory board game, called "Data Iceberg", helps players learn about data by recognising and categorising different types of data in various everyday situations. The game

uses a deck of cards and group boards with rectangular spaces to place pairs of cards. In each pair, there is a statement describing the type of data shown in the picture. An example of a pair of cards and part of the card holder board is shown in Figure 2. Each rectangular space represents a data category. Players work in teams, with teamwork and discussion strongly encouraged.
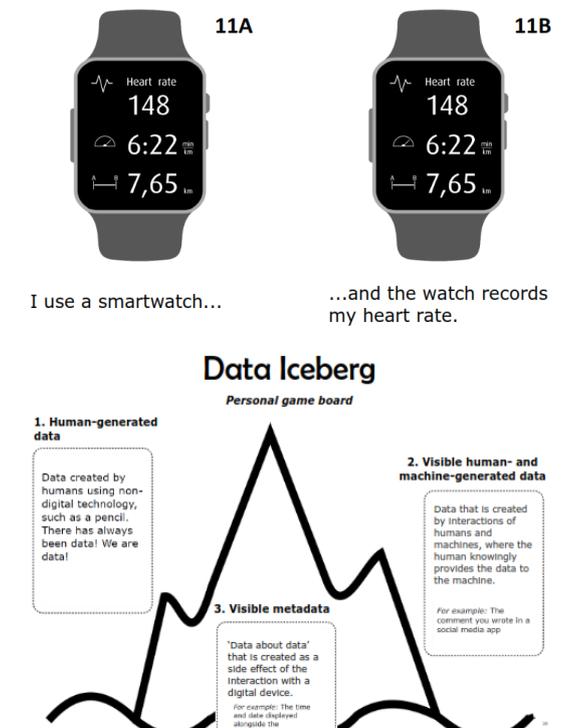
Figure 2: Example of the memory cards and a part of the card holder board
(DALI Project, 2023b)

After the game, the facilitators recorded how many pairs of cards each group correctly placed in the rectangles representing the types of data described on the cards. The facilitator then explained the correct placement of all pairs of cards and the rationale for their placement. Group scores were announced, and no further analysis of the data was conducted.

Post-test
The main purpose of the post-test was to assess changes in participants' understanding of data types after the workshop. The post-test was also intended to evaluate the workshop's impact and effectiveness. All workshop participants completed the post-test at the end of the workshop. The pre- and post-test instruments were identical, as shown in Appendix 1. After all participants completed the post-test, a feedback session was held to discuss the instrument items and explain the correct answers. Once all pre- and post-test data had been recorded, we checked the data for normality and identified outliers. The data were normally distributed and contained no outliers. We also ensured independence between pre- and post-test scores by administering the pre-test two weeks before the post-test. The post-test was administered immediately after the workshop ended. We believed this arrangement ensured the independence of the pre- and post-test scores. After all assumptions had been

met, we applied descriptive and inferential (paired t-test) statistical analyses to assess whether there were any changes in participants' understanding of data types after the workshop. Following the workshop, interviews were conducted to gather in-depth qualitative feedback from participants about their experiences, perceptions, and learning outcomes.

Interviews

The main purpose of the interviews was to deepen, explain, and contextualise the pre- and post-test results. The interviews aimed to clarify why some participants improved more or less than others and to capture participants' personal perspectives on the workshop's learning activities. Three participants were selected for interview after taking the post-test. They were chosen based on their scores on the pre- and post-tests: the most improved, one with no improvement, and one with a learning decline. All three participants enjoy volunteering, adventure, and nature.

- The most improved interviewee (R1) is a 32-year-old female facilitator and founder of her TBM, whose post-test score increased the most (by 5 points). She is very active in the literacy movement in Aceh. She was born and raised in a small district in Aceh Province, where she established the TBM. She has extensive experience as an educator at the national level and holds a bachelor's degree in the humanities from a state university in Aceh.
- The no improvement interviewee (R2) is a 23-year-old male facilitator who is a recent Management graduate from a state university in Aceh. He was a new volunteer at his TBM. He answered the same questions correctly in both the pre- and post-tests.
- The interviewee who experienced a learning decline (R3) is a 29-year-old female facilitator who has worked at her TBM for more than five years. She has a bachelor's degree in humanities from a state university in Aceh, Indonesia. Her pre-test score was 2 points higher than her post-test score.

The instrument used in the interview was a semi-structured interview, with questions presented in Appendix 4. The interviews were audio-recorded and transcribed verbatim. Identifying information was removed to ensure confidentiality and anonymity. The transcripts were then systematically coded manually using an inductive approach. Related codes were grouped into categories, which were further refined into themes representing recurring patterns across interviews.

Observations

The main purpose of the observations conducted during the workshop was to capture how participants engaged with the workshop activities, materials, and facilitators. Using an Observation Rubric (see Appendix 5), researchers systematically observed, listened to, and recorded participants' engagement, interactions and collaboration, use of workshop materials, and facilitator-participant interactions. This provided data on how participants engaged with the activities, their responses, and the dynamics of the learning process. Each indicator was rated on a three-point scale. The data were then triangulated with pre- and post-test results and follow-up interviews.

Ethics statement

This study was conducted in accordance with recognised ethical standards for educational inquiry. Participation was voluntary, and all participants were informed of the study's purpose, the procedures involved, and their right to withdraw at any time without repercussions. All participants provided informed consent before data collection commenced. The anonymity and confidentiality of participants were ensured by removing

identifying information and securely storing all data. Audio recordings and texts were used exclusively for research purposes and were accessible primarily to the research team. No harm to participants was expected or experienced as a result of their involvement in the study.

The authors used OpenAI's ChatGPT version 5 and Quillbot to enhance language and readability during the preparation of this manuscript. ChatGPT was also used to initiate the literature review process by finding robust references. An example of a prompt used for ChatGPT is: "Can you provide references on mixed-method research design?" The authors used Quillbot for paraphrasing to ensure a concise and academic writing style. Following the use of these language models, the authors reviewed and made necessary edits to the content, assuming full responsibility for the publication's content. ChatGPT and Quillbot cannot be listed as authors, and the authors take full responsibility for the accuracy and integrity of all content.

RESULTS

Demographics of the participants
The participants in this study are members and facilitators of Forum TBM Aceh. All invited TBMs were selected from the list of Forum TBM Aceh members, located in Banda Aceh City and Aceh Besar District. Thirty-five of the 47 invited TBM facilitators participated in this study, comprising 12 males and 23 females. Participation in this study was by invitation, and no random selection was used, as we intended to include all TBM facilitators in Banda Aceh City and Aceh Besar District in the FGD and workshop. Facilitators who could benefit from the intervention may be denied access to potentially significant educational opportunities if they are excluded.

Participants' initial data literacy levels, in terms of understanding data (pre-test results)
From the FGD, we found that all participants had frequently used data in their community libraries' activities; however, they were not aware of the extent of data involvement in these activities. They only realised they were producing data from their posts, such as comments, and from all their social media activities. They did not realise that when they reported an activity, they were working with data, and that when they made decisions, they sometimes used data as the basis. We found that most participants were at the basic level of the DALI Framework's components (understanding data, acting on data, and engaging through data).

Table 1 presents a summary of descriptive statistics derived from participants' responses to the pre-test and post-test, comprising 10 multiple-choice questions designed to assess their capacity to differentiate data types based on their origin (human, machine, or a combination) and their nature (raw data or metadata, visible or invisible). All questions had the same answer options. If a participant answered an item correctly, he or she received one point. For example, if a participant answered 5 items correctly, the total score would be 5 out of 10. The questionnaire was administered via a Google Form, with an answer key provided, so the total score was recorded directly in the response table.

Table 1: Summary of descriptive statistics of pre-test and post-test scores

	Pre-Test	Post-Test
Participants (N)	35	35
Mean Score	3.60	4.69
Variance	3.01	4.58
Standard Deviation	1.74	2.14

The mean pre-test score is 3.60 with a standard deviation of 1.74, indicating that participants, on average, scored significantly below the midpoint, which suggests limited understanding of data types prior to the workshop. The standard deviation of 1.74 shows moderate variability in scores and noticeable differences among participants. While some participants scored slightly higher, others performed considerably lower than the group average.
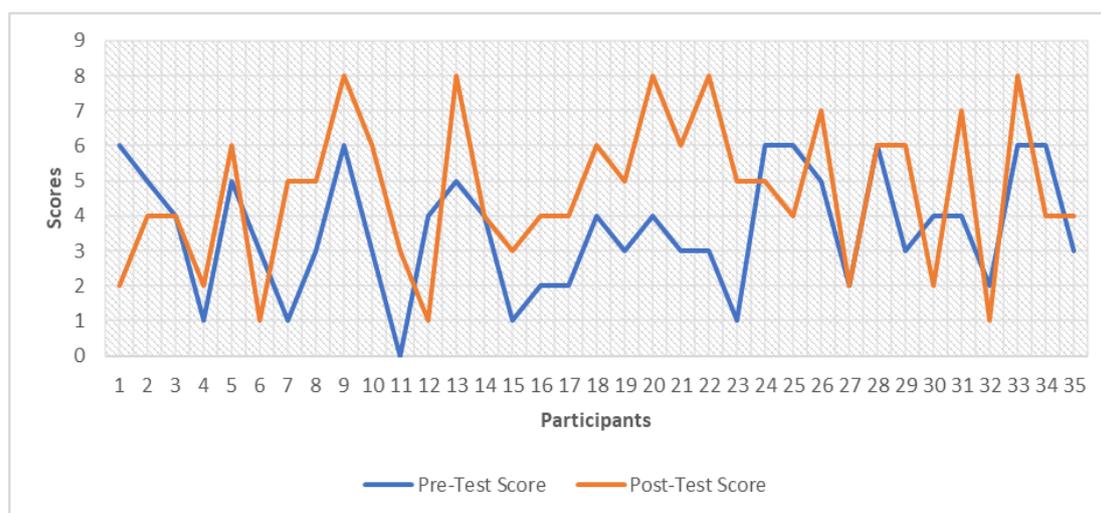

Figure 1: Pre-test and post-test scores of 35 participants

Participants' data literacy improvement in understanding data (post-test results)
Figure 3 shows the comparison of each participant's pre-test and post-test scores. Some participants scored lower in the post-test than in the pre-test, suggesting that certain individuals experienced a decline in their comprehension of data types following the workshop.

The post-test results show a mean score of 4.69 with a standard deviation of 2.14, indicating a moderate but still below-average level of performance. This suggests that participants demonstrated only partial or emerging understanding of data types. The moderate increase in mean test scores from 3.60 in the pre-test to 4.69 (out of 10) in the post-test suggests that the workshop contributed to some improvement in participants' understanding of data types. The null hypothesis stated that the means of the two test score groups are equal; a paired t-test produced a p-value below the significance threshold of 0.05. This result indicates a statistically significant difference in the mean scores of the two groups, with post-test scores, on average, only 1 point higher than pre-test scores.

The magnitude of the effect can be quantified with Cohen's d formula (Cohen, 2013). The calculated effect size, d = 0.487, shows that the improvement exceeded the threshold for a small effect ($|d| = 0.2$) but was slightly below the threshold for a medium effect ($|d| = 0.5$),

and remained well below the threshold for a large effect ($|d| = 0.8$). This suggests that the workshop had a moderate, meaningful effect on participants' understanding of data types. The improvement is noticeable and practically relevant, though further activities may be needed to achieve a stronger effect.

Participants perceived learning experience

Participants generally reported a positive learning experience during the DALI-based workshop, although there were minor discrepancies in depth and confidence. The details are provided below:

i.   *Content relevance*

Some participants found the workshop content relevant to their needs, particularly in understanding types of data, while others noted that the content could be improved. For example, one of the memory cards used in the board game contains statements about data transmitted by an electric car to its manufacturer, enabling the manufacturer to assess the driver's driving style accurately, as shown in Figure 4. This context is irrelevant for the participants, as not all are familiar with advanced cars equipped with technology capable of transmitting such complex data. We attempted to simplify the task by changing the statement "the manufacturer knows exactly the driving style of the driver" to "the manufacturer knows exactly the route taken by the car." Although some participants are familiar with this technology, most are unable to determine whether the data transmitted by the vehicle qualifies as raw data or metadata. This is overly complex for those who have never encountered such advanced technology in a vehicle. We acknowledge that we implemented the DALI Framework without fully considering participants' access levels.
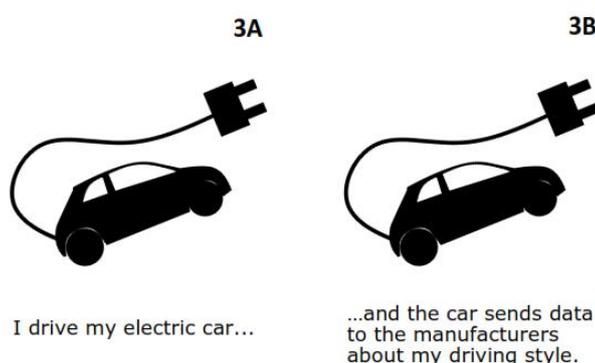

Figure 4: Memory cards with identical images and explanations about the category of data they represent (DALI Project, 2023b)

Despite the lack of increased knowledge, participants demonstrated greater awareness of data use and production, particularly through the mini-lecture and memory board game. The most improved participant noted, "The workshop increased my awareness of the algorithm linking my social media platform to external sources based on my activities on the platform (R1)." She reflected, "We are data creators, not just data consumers (R1)," while the participant whose understanding declined commented:

"I liked the first part of the game, but we got lost in the second part when we tried to group the types of data that match the data on the cards. When the workshop and the post-test were over, I finally understood how different pieces of data are based on their traits" (R3).

These statements capture a critical shift in perception: recognising everyday interactions with data and the associated power dynamics of data ownership.

The transition from workshop knowledge to community application proved a significant challenge, despite positive engagement and learning. Participants identified limited digital infrastructure and insufficient community knowledge of data issues as obstacles. The most improved participant stated:

"If we are given a training of trainers on the game application and have resources in Indonesian that are suitable for our community's age group, I am sure I could use the game at my TBM. It is good for our group to learn how to use data well" (R1).

The interviewee whose understanding regressed suggested using examples of relevant data types for their TBM activities in the memory board game. To enhance data literacy through the DALI Framework, it is essential to localise the framework by incorporating local language and examples. Common indicators such as market prices, weather patterns, school attendance, book genres, and crop yields are of greater significance to the general population in developing countries. Furthermore, it is crucial to simplify pathways for low-literacy populations and develop strategies to engage underserved groups.

ii. Instructional clarity
All three interviewees expressed a preference for the interactive instructional methods used in the workshop. Facilitators were regarded as clear and supportive. Observations from the workshop showed significant engagement during group activities, especially the memory game board activity. Engagement was characterised by frequent exchanges of ideas and clarifying questions among participants. Although some participants were initially hesitant, they became increasingly vocal as the session progressed, indicating a supportive learning environment. After the workshop, their understanding of certain abstract data concepts improved. The differences between raw data and analytical data, as well as metadata and its creation and use, are now understood.

iii. Confidence development
The most improved participant reported increased confidence in her ability to interpret and engage with data following the workshop, as well as a readiness to undertake similar activities at her TBM. She stated:

"The game is a lot of fun, and it teaches us about different kinds of data. We are data creators, not just data consumers. Every day, we make data, and anyone can access it without our permission" (R1).

The participant experienced a learning decline and demonstrated confusion during the workshop and in the post-test. However, she regained her confidence after the differences between data types were summarised following the post-test, as stated below:

"I thought all the data that computers make is analytical data, metadata, not raw data. However, after the wrap-up just now, the differences are very clear to me. I should have known that before the test, but I did not" (R3).

The most improved participant demonstrated increased confidence and readiness to use her data knowledge, whereas the participant who experienced a learning decline showed reduced confidence when answering the post-test. She mentioned:

"I used my intuition to answer the pre-test, but after learning about the different types of data, I got lost and did not trust my intuition to answer the post-test" (R3).

This highlights the need for differentiated instruction and iterative reinforcement in future workshops. Engagement levels significantly influenced learning outcomes. Participants who engaged collaboratively, asked clarifying questions, and gradually built confidence, especially during the workshop's interactive elements, showed greater improvement. This finding aligns with adult learning theories that emphasise experience and social learning as essential for meaningful knowledge construction (Knowles et al., 2025; Brookfield, 2013).

iv. Applicability to real-world tasks

Some participants were able to connect workshop activities to real-life situations, recognising how data literacy skills could inform decision-making, project planning, and professional tasks. Participants acknowledged the significance of data literacy in their professional responsibilities and expressed a desire to incorporate data-driven methodologies into their TBM activities. All interviewees demonstrated a keen interest in disseminating their new knowledge of data literacy among colleagues and in enhancing organisational procedures through improved data use, thereby significantly improving efficiency, transparency, and impact.

Following the workshop, several requests were made for additional practical sessions. Interviewees suggested incorporating real-world case studies from literacy communities to improve relevance. When asked about their views on the proposed sharing of resources among TBMs linked to Forum TBM Aceh, all interviewees agreed, indicating that they see this as a beneficial approach.

DISCUSSION

The results showed that, prior to the workshop, participants' understanding of data types was limited. The workshop's learning activities helped some participants improve their understanding. This was evident from the statistical analysis conducted after the post-test scores were collected. Descriptive analysis revealed that participants' post-test scores were, on average, one point higher than their pre-test scores. Despite the statistical significance of this improvement, as confirmed by the paired t-test, the increase in mean test scores from 3.60 to 4.69 out of 10 is moderate. In general, participants could not answer 50% of the test items. These findings suggest that the workshop did enhance participants' performance on identical test questions after they received training on data types and their properties; nonetheless, the improvement is modest. Consequently, the training only slightly enhanced participants' understanding of data types and their characteristics. Furthermore, the effect size observed was just below the threshold for a medium effect, indicating that while the training had a positive impact, it was modest in magnitude.

Several factors have been identified as potentially contributing to this outcome. The DALI Framework, first and foremost, aims to support digital education for European citizens, assuming the regular use of digital tools, databases, and online information sources. Even though Castañeda et al. (2024) suggested that the DALI Framework's "adaptability and scalability make it well-suited for addressing the diverse needs and contexts found within different educational settings and organisations internationally (p. 1)," careful planning and adaptation of the learning materials to be more relevant for participants from developing countries should be considered. In this study, not every TBM facilitator consistently uses

databases, online information sources, and digital technologies. For them, certain digital terms such as metadata and raw data are rather new. Furthermore, in some regions of Aceh, electricity and internet connectivity are unreliable, making it difficult for people to practise and apply the skills required by the DALI framework.

Another factor likely contributing to the limited improvement in participants' understanding of the data is the quality of the translations and the cultural contexts in the workshop's learning materials. The absence of immediate utility within their TBM contexts was perceived not as a rejection but as a temporary gap, acknowledging that as digital technologies advance, the imperative to integrate data literacy will likewise intensify. This proactive stance highlights participants' awareness of emerging needs despite current structural limitations and indicates the possibility of gradual integration of data practices. With continued exposure to and access to the training resources, participants are expected to further develop their data literacy skills in alignment with the DALI Data Literacy Framework, particularly as they apply these skills in their TBM activities.

The programme fostered a positive attitude towards data utilisation, with participants recognising its capacity to optimise operations, improve transparency, and increase impact within their TBMs. Participants demonstrated both learning transfer and an increasing ability to conceptualise and propose system-level improvements based on data logic. This highlights the motivational significance of perceived value in learning retention and future behaviour change. Participants provided constructive recommendations for improving future training sessions, including incorporating real-world case studies and TBM-specific data examples. These recommendations indicate a need for greater contextual relevance and authenticity, which is particularly important for adult learners in an informal educational environment.

In general, participants perceived their learning experience during the DALI-based workshop as meaningful and beneficial. Although some participants suggested using materials more relevant to community library tasks, others considered that the workshop facilitator delivered the content clearly, especially after the workshop concluded. The wrap-up session at the end of the workshop was the key component that helped participants clarify their understanding. Most participants reported increased confidence after the workshop and were ready to apply similar tasks in their community library activities.

CONCLUSIONS

Overall, the workshop provided meaningful exposure to key data concepts and valuable opportunities for participants to strengthen their foundational understanding of data types and their properties. The observed gains from the pre-test, though incremental, indicate measurable learning and affirm the workshop's role in supporting the development of data literacy. These initial improvements establish a strong baseline for continued growth. With strategic planning, thoughtful adaptation of the DALI Framework, and the use of contextually relevant training materials, future iterations of the workshop are well positioned to further enhance learning outcomes, enabling participants to deepen their data literacy skills and apply them more confidently and effectively in subsequent activities.

The findings suggest that data literacy training, when grounded in participatory and contextualised approaches, can significantly enhance TBM facilitators' awareness of data use and production, as well as build confidence and motivation to help their community

become data literate. Despite ongoing implementation challenges, particularly in low-resource settings, the potential for long-term transformation is evident. Future initiatives must emphasise sustainable support, contextually relevant content, and practical applications that align with participants' daily activities. TBM can use data not only for reporting and compliance but also as a strategic asset for community empowerment and organisational growth.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

The authors have no relevant competing interests to declare pertaining to the content of this article.

AUTHOR CONTRIBUTIONS

Conceptualisation: [Oktavia R. & Yuni S.M.] Methodology: [Oktavia R. & Yuni S.M.], Formal analysis and investigation: [all authors], Writing-original draft preparation: [Oktavia R.] Writing-review and editing: [all authors].

REFERENCES

Bowler, L., & Shaw, C. (2024). Trends in data literacy, 2018-2023: a review of the literature. *Information Research an International Electronic Journal, 29*(2), 198–205. https://doi.org/10.47989/ir292822

Brookfield, S. (2013). Teaching for Critical Thinking. *International Journal of Adult Vocational Education and Technology, 4*(1), 1–15. https://doi.org/10.4018/javet.2013010101

Calzada Prado, J., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri, 63*(2), 123–134. https://doi.org/10.1515/libri-2013-0010

Castañeda, L., Haba-Ortuño, I., Villar-Onrubia, D., Marín, V. I., Tur, G., Ruipérez, J. A., & Wasson, B. (2024). Developing the DALI Data Literacy Framework for critical citizenry. *RIED-Revista Iberoamericana de Educacion a Distancia, 27*(1), 289–318. https://doi.org/10.5944/ried.27.1.37773

Cohen, J. (2013). Statistical power analysis for the behavioral sciences. routledge.

Cui, Y., Chen, F., Lutsyk, A., Leighton, J. P., & Cutumisu, M. (2023). Data literacy assessments: a systematic literature review. *Assessment in Education: Principles, Policy and Practice 30*(1), pp. 76–96. Routledge. https://doi.org/10.1080/0969594X.2023.2182737

DALI Project. (2023a). *DALI data literacy framework*. https://dalicitizens.eu/index.php/dali-data-literacy-framework-2/

DALI Project. (2023b). *Data iceberg* [Board game]. https://toolkit.dalicitizens.eu/catalogue

DALI Project. (2023c). *DALI facilitators guide*. https://toolkit.dalicitizens.eu/for-facilitators

DALI Project. (2023d). *Game based networked learning system handbook*. https://toolkit.dalicitizens.eu/for-facilitators

Fontichiaro, K., & Oehrli, J. A. (2016). Why Data Literacy Matters. *Knowledge Quest, 44*(5), 21–27. https://files.eric.ed.gov/fulltext/EJ1099487.pdf

Forum TBM. (2024). *Forum TBM: Selayang Pandang.* https://forumtbm.or.id/tentang-kami/

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

Gummer, E. S., & Mandinach, E. B. (2015). Building a conceptual framework for data literacy. *Teachers College Record*, *117*(4), 1–22.

Knowles, M., Holton III, E.F., Robinson, P.A., & Caraccioli, C. (2025). *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development* (10th ed.). Routledge. https://doi.org/10.4324/9781003387671

Koltay, T. (2017). Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science, 49*(1), 3–14. https://doi.org/10.1177/0961000615616450

Lee, J., Alonzo, D., Beswick, K., Oo, C. Z., Anson, D.W.J., & Abril, J.M.V. (2024). Data literacy of principals in K–12 school contexts: A systematic review, *Educational Research Review, 45*,100649, https://doi.org/10.1016/j.edurev.2024.100649.

Mandinach, E. B., & Gummer, E. S. (2016). *Data literacy for educators: Making it count in teacher preparation and practice*. Teachers College Press.

Matuk, C., DesPortes, K., Amato, A., Silander, M., Vacca, R., Vasudevan, V., & Woods, P. J. (2021). Challenges and Opportunities in Teaching and Learning Data Literacy through Art. In de Vries, E., Hod, Y., & Ahn, J. (Eds.), *Proceedings of the 15th International Conference of the Learning Sciences - ICLS 2021*. (pp. 681–684). Bochum, Germany: International Society of the Learning Sciences. https://repository.isls.org//handle/1/7556

Molina-Azorin, J. F. (2016). Mixed methods research: An opportunity to improve our studies and our research skills. *European Journal of Management and Business Economics, 25*(2), 37–38. https://doi.org/10.1016/j.redeen.2016.05.001

Novrita, J., Elizarni, Oktavia, R., & Sari, T. Y. (2025). Making 'Taman Baca' Sustainable", lessons learned from community-based non-formal education in Aceh, Indonesia. *International Journal of Educational Development*, 113. https://doi.org/10.1016/j.ijedudev.2024.103186

Nwagwu, W.E. (2025). Citation analysis of global research on data literacy. *Educ Inf Technol 30*, 15931–15959. https://doi.org/10.1007/s10639-025-13397-9

OECD. (2024). PISA 2022 Results: Volume III: Creative Minds, Creative Schools Factsheets – Indonesia. https://www.oecd.org/en/publications/pisa-results-2022-volume-iii-factsheets_041a90f1-en/indonesia_a7090b49-en.html

Ongena, G. (2023). Data literacy for improving governmental performance: A competence-based approach and multidimensional operationalization. *Digital Business, 3*(1), 100050. https://doi.org/https://doi.org/10.1016/j.digbus.2022.100050

Pinto, M., Caballero-Mariscal, D., García, F. J. M., & Gómez-Camarero, C. (2023). A strategic approach to information literacy: data literacy. A systematic review. *Profesional de La Informacion, 32*(6). https://doi.org/10.3145/epi.2023.nov.09

Pothier, W., & Condon, P. (2025). Data Literacy Skills: Industry Perspectives and Professional Practice. *Portal: Libraries and the Academy, 25*(2), 271–298. https://doi.org/10.1353/pla.2025.a955946

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., & Wuetherick, B. (2015). *Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report Key Messages*. http://hdl.handle.net/10222/64578

Rubin, A., & Gould, R. (2023). Introduction. In G. F. Burrill, L. de Oliveria Souza, & E. Reston (Eds.), *Research on Reasoning with Data and Statistical Thinking: International Perspectives* (pp. 1–8). Springer International Publishing. https://doi.org/10.1007/978-3-031-29459-4_1

Schreiter, S., Friedrich, A., Fuhr, H., Malone, S., Brünken, R., Kuhn, J., & Vogel, M. (2024). Teaching for statistical and data literacy in K-12 STEM education: a systematic review on teacher variables, teacher education, and impacts on classroom practice. *ZDM - Mathematics Education, 56*(1), 31–45. https://doi.org/10.1007/s11858-023-01531-1

Stokes, W. T. (2008). *Literacy In Encyclopedia of Educational Psychology, Vol 2*, pp. 609–616. https://doi.org/https://doi.org/10.4135/9781412963848.n161

Tashakkori, A., Johnson, R. B., & Teddlie, C. B. (2020). *Foundations of Mixed Methods Research. Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, (2nd ed). Sage Publication, US.

The World Bank. (2023). *Literacy rate, adult total (% of people ages 15 and above) - Indonesia*. https://data.worldbank.org/indicator/SE.ADT.LITR.ZS?end=2020&locations=ID&start=2014

UNESCO. (2019a). *UNESCO strategy for youth and adult literacy (2020-2025).* https://unesdoc.unesco.org/ark:/48223/pf0000371411.locale=en

UNESCO. (2019b). *Advancing lifelong learning through Community Learning Centers: Guidance for policymakers — Asia and the Pacific*. UNESCO Bangkok.

UNESCO. (2025, January 31). *What you need to know about literacy*. https://www.unesco.org/en/literacy/need-know

UNESCO Institute for Information Technologies in Education (IITE). (2021). *Digital literacy in education: Policy brief*. UNESCO.

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).

Appendix 1: Pre-test and post-test items

These are ten translated items from the pre-test and post-test. Participants may select one of five options for each question, with only one correct answer. The items are listed below.

No	Question	Scale
1	What type of data is your tweet on the X platform (formerly Twitter) about your disappointment with the refereeing in the World Cup Qualifier football match between Indonesia and Bahrain?	▪ Visible Analytical Data ▪ Visible Raw Data ▪ Invisible Analytical Data ▪ Invisible Raw Data ▪ Functional Data
2	What type of data is your tweet on the X platform (formerly Twitter) about your disappointment with the refereeing in the World Cup Qualifier football match between Indonesia and Bahrain?	
3	To plan my next bus trip, I use an app where I add my bus stops and destination. The app shows me which buses I can take. What type of data are the bus stops and travel destinations I added?	
4	I visited a website and clicked on several links on that website. What type of data is generated by clicking those links?	
5	I drive my electric car, and it sends data to the manufacturer about my driving style. The data sent by the car to the manufacturer is…	
6	The Internet Protocol (IP) enables communication and data exchange between devices on the Internet. IP specifies the address of the device intended to receive the data. To what type of data does the Internet Protocol (IP) relate?	
7	I use a smartwatch to record various biomarkers. For example, the watch records my heart rate and displays the different heart rate zones I experience during exercise. What type of data is the visualisation generated by the smartwatch?	
8	To listen to music, I use the Spotify app where I created a profile from scratch, without using my social media accounts. What type of data is the profile I created from scratch?	
9	To prepare for my next holiday, I used my search engine to find a good road route in the area. My search engine saved the websites I visited, as well as my clicks and mouse movements on those websites. What type of data is stored by the search engine?	
10	I booked a flight using my laptop and received ads for hotels in the same destination on my phone's browser. What type of data are the ads that appear in my phone's browser?	

Appendix 2: Semi-structured focus group discussion framework

Participants: Facilitators of Community Libraries
Subject: Experiences with data and data literacy in community library environments

---
Introduction / Preliminary Activity
---

1. Please provide a brief introduction of yourself and outline your responsibilities within the community library.
    Inquiry: How long have you served as a library facilitator?

2. In your role, how do you typically engage with or use data in managing library programmes or services?

Section 1: Encounters with Data
---

3. Could you describe your experiences using data in your community library?
4. How do you typically use data to plan, improve, or evaluate library activities or services?

Probe: Decision-making, reporting, funding applications, or community engagement

5. How confident do you feel when handling data related to library operations or programmes? Why?

Section 2: Categories of Data Encountered
---

6. What types of data are most frequently available or collected at your library?
7. Which categories of data do you find most manageable, and which present greater challenges?
8. How do you typically acquire or obtain data in your library environment?

Section 3: Comprehending and Utilizing Data
---

9. How do you interpret the data you gather or obtain?
Probe: Using summaries, charts, consultations with colleagues, or personal assessment.

10. What challenges do you encounter when analysing data or determining its application?
Investigation: Insufficient skills, ambiguous data, inadequate instruments, or time limitations.

Appendix 3: Instrument of memory board game

1A Saya pergi berjalan lintas alam...	**1B** ...dan saya menggunakan ponsel saya untuk merekam rute perjalanan saya.
2A 08.12.2021 11:51 Saya menyukai sebuah foto teman saya di media sosial	**2B** 08.12.2021 11:51 ...dan saya dapat melihat data mengenai waktu saya menyukai foto tersebut.
3A Saya mengendarai mobil listrik saya...	**3B** ...dan mobil tersebut mengirim data ke produsen mobil tentang rute yang saya tempuh.
4A Saya menghitung semua perjalanan saya setiap bulan...	**4B** ...dan saya menampilkannya dalam bentuk grafik di buku tulis dengan menggunakan pena.
21A Saya bermain game online...	**21B** ...dan saya memberi informasi mengenai umur saya.
22A Saya menggunakan app untuk mempelajari bahasa asing...	**22B** ...dan app tersebut menunjukkan data perkembangan belajar saya
23A Saya menggunakan laptop untuk mencari mainan baru...	**23B** ...dan saya mendapatkan iklan mainan di browser ponsel saya.
24A Saya menggunakan app filter wajah...	**24B** ...dan app tersebut mengumpulkan data lokasi saya tanpa menunjukkannya

Data Iceberg
Personal game board

1. Data yang dihasilkan manusia

Data created by humans using non-digital technology, such as a pencil. There has always been data! We are data!

2. Data yang terlihat yang dihasilkan oleh manusia dan mesin

Data that is created by interactions of humans and machines, where the human knowingly provides the data to the machine.

For example: The comment you wrote in a social media app

3. Metadata yang terlihat

'Data about data' that is created as a side effect of the interaction with a digital device.

For example: The time and date displayed alongside the comment you posted in the social media app.

4. Metadata tersembunyi

'Data-about-data' that is created by machines based on extracted information from a users activities. This is done in a non-transparent manner, which means it is done without the awareness of the human subject (you).

For example: The order in which you visit specific websites, and the time you spent on it, is collected through a common web-statistics service.

DaLi 010110I
Data Literacy for Citizenship

Appendix 4: Interview questions

1. Please provide an overview of your experience during the workshop. What stood out to you most?
2. Could you give examples of specific moments or activities from the workshop that you found especially engaging or challenging?
3. How did the workshop improve your understanding of types of data?
4. What challenges might arise if the approach used in this workshop is implemented to teach data literacy in your TBM?

Appendix 5: Observation rubric

Dimension	Indicators	1 – Low	2 – Moderate	3 – High
Engagement with Data Activities	Attention, participation in hands-on data tasks	Rarely participates; disengaged from data tasks	Participates inconsistently; needs prompting	Actively and consistently participates
Understanding of Data Concepts	Data terms, interpretation, reasoning	Shows limited or incorrect understanding	Demonstrates partial understanding	Demonstrates clear and accurate understanding
Use of Data Tools and Resources	Worksheets, datasets, digital tools	Unable or reluctant to use tools	Uses tools with guidance	Uses tools independently and effectively
Data Interpretation Skills	Reading charts, tables, datasets	Misinterprets or avoids data	Basic or partially correct interpretations	Accurate and confident interpretation
Critical Thinking with Data	Questioning data sources, assumptions	Rarely questions data	Occasionally questions data	Consistently evaluates data critically
Collaboration and Communication	Discussion, explanation, peer support	Minimal interaction	Some collaborative discussion	Active discussion and peer support
Application to Real-Life Contexts	Linking data to real-world issues	Unable to make connections	Makes general connections	Clearly applies data to real contexts
Facilitator: Participant Interaction	Asking questions, responding to feedback	Rarely interacts	Occasional interaction	Frequent, meaningful interaction