The profile of students’ data literacy at SMA Negeri 1 Karanganyar

Aniks Ambarwati\(^{(1,a)}\), Sarwanto\(^{(2,b)}\), Sukarmin\(^{(1,c)}\)

\(^{1,2}\)Physics Education, Master Program, Universitas Sebelas Maret, Surakarta, INDONESIA
\(^{1}\)Science Education, Master Program, Universitas Sebelas Maret, Surakarta, INDONESIA

Email: \(^{a)}\)ambarwatianiks@student.uns.ac.id; \(^{b)}\)sarwanto@fkip.uns.ac.id; \(^{c)}\)sukarmin67@staff.uns.ac.id

Abstract. The new 4.0 industry has a significant impact on the 21\(^{st}\)-century advanced development. The development of science and technology has experienced rapid changes and progress. Students must prepare themselves with various competencies to face the challenges that arise in global competition. One of the competencies that must be owned by the 21st-century skills of students is data literacy. Data literacy is the ability of individuals to collect, process, evaluate, and apply data. The purpose of this study was to find out the literacy profile of senior high school students, especially in physics learning. The data collection technique in this study was to provide tests to several students in SMA Negeri 1 Karanganyar. The data obtained analyzed using descriptive statistical analysis. Based on the tests results obtained an average score of high school student’s data literacy 55.72%. The results received described the student data literacy profile, in general, were still in the medium category.

1. Introduction

During the rapid development of technology and information in the 21st century, education has many opportunities and challenges in every aspect of life. The quality of human as a social creature determines how they can take advantage of these opportunities and challenges to adapt and develop existence with the development of the times. One effort in improving the quality of human resources is through education. Formal and non-formal education nowadays is the right target to create various competencies, and skills that are useful for students to face global competition. One of these competencies related to global challenge is the ability to collect, process, evaluate and apply data called data literacy.

Education is the beginning in the step of teaching and learning process. It needs to prepare their graduates with more data-based skills for the workforce and the community as a whole, which is increasing centered on data [1]. Providing graduates with basic knowledge of data literacy makes it possible for students entering the 21st century workforce to apply various skills to a variety of situations. Industry feedback has also provided insight into what skills valued in the workforce, and extraordinary responses are focused on data-related skills [2]. Data literacy is an essential capability demanded in economic-based global knowledge, and data manipulation activities are carried out in everyday processes across sectors and fields. An understanding of how to make decisions from data that is informed and how to collect, process, evaluate, and apply data in support of facts based on
decision making [3]. By developing data literacy, students also sharpen their skills in making decisions, answering a question by properly using data, interpreting your findings, and receiving action information [4].

Learning at schools is a medium in developing data literacy for students, one of them through learning physics. Physics learning is not only focused on mastering concepts alone but also as a vehicle to foster thinking skills that are useful for solving problems in everyday life. Physics learning in schools also equips students with the knowledge, planting attitudes and skills that are useful when entering higher levels of education and challenges in the world of work. Based on the description above, the formulation of the problem arises, namely how is the data literacy profile of high school students, especially in physics learning?

1.1. Data Literacy

According to the Australian Public Service Commission [5], the term data literacy refers to the ability to identify, find, interpret and ethically evaluate information and then use that information to answer or respond to specific questions or problems. This data literacy is one of the crucial aspects which must be embedded in learning. The development of data literacy capabilities it also generates other abilities such as the ability to think critically, analytically, logically, and others. Someone capable of data literacy has the knowledge to gather, analyze, and graphically convey information and data to support decision making [6]. Individuals who have literacy skills can understand, explain, and document the usefulness and limitations of data. This can been done by becoming consumers of critical data, controlling the traces of their data, finding meaning in data, and taking actions based on data. Data literate individuals to be able identify, collect, evaluate, analyze, interpret, present, and protect data [7].

Data literacy in conceptual term requires critical thinking skills, gaining knowledge from abstraction and the application of results [8,9]. Critical and abstract thinking skills are similar to computational thinking. Computational thinking involves the ability to define abstractions, work with several layers of abstraction and understand relationships between different layers [10]. Data itself is an abstraction and be used if the student can understand how it relates to other information on the broader world. This ability to understand intricate relationships and connections further increases the capacity for curiosity and in depth thinking [11].

1.2. Competence in Data Literacy Framework

A study done showed that data literacy competencies include the ability to understand, search, collect, interpret, visualize, and support arguments using qualitative and quantitative data [12]. Several data literacy components can be distinguished: set a purpose, collect data, analyze data, interpret data, and take instructional action [19]. According to the School of Data in an online lecture entitled Data Fundamentals, it explained that data literacy covers six topics, among others: understanding data, searching for data, sorting and filtering data, analyzing data, making data visualization and telling stories using data. Whereas according to Carlson et al., there are twelve main fields of data literacy competencies, namely introduction to database and data format, data discovery and acquisition, data management and organization, data conversion and separability, quality assurance, metadata, data curation and reuse, practice culture, data preservation, data analysis, data visualization, and ethics [13]. Furthermore, data literacy competencies according to Javier et al. can be explained as follows [14]:

1.2.1. Understanding Data.

Data literacy requires an understanding of what data is, what types of data exist, and how data is generated. A data literate individual must also have knowledge of the role and impact of data in society across different contexts and the ethical implications of using data.
1.2.2. Finding and/or obtaining Data.
Data literacy requires of the possible the data sources, be able to evaluate them and select the ones most relevant to an informational need or a given problem. Learners need to be able to detect when a given problem or need cannot be solved with the existing data and, as appropriate, undertake research to obtain new data.

1.2.3. Reading, interpreting and evaluating data.
Data literacy required to be aware of the various forms in which data can be presented and be able to interpret them. Learners need to be able to evaluate data critically.

1.2.4. Managing Data.
Data literacy requires the ability to save the data selected or generated and of descriptive or other data associated therewith, for due identification, management and subsequent reuse.

1.2.5. Using Data.
Data literacy includes the ability to prepare data for analysis, analyse them in keeping with the results sought and know how to use the necessary tools. Learners need to be able to synthesize and represent the results of data analysis in ways suited to the nature of the data, their purpose and the audience targeted in the inquiry. Data literacy includes the ethical use of data, acknowledging the source and make sure that used methods are deployed and results interpreted transparently and honestly.

The competence of data literacy framework are divided into two essential parts, namely context and strategic values. It is related to data literacy education and best practices for teaching data literacy in every discipline. More specific data literacy competencies is divided into several steps. Each step has its skills, as explained through the flowchart in Figure 1 [3].

![Conceptual Framework](chart)

**Data Collection**
- Data Discovery and Collection
- Evaluating and Ensuring Quality of Data and Source

**Data Management**
- Data Organization
- Data Manipulation
- Data Conversion (from format to format)
- Metadata Creation and Use
- Data Curation, Security, and Re-Use
- Data Preservation

**Data Application**
- Critical Thinking
- Data Culture
- Data Ethics
- Data Citation
- Data Sharing
- Evaluating Decision Based on Data

**Data Evaluation**
- Basic Data Analysis
- Data Interpretation (Understanding Data)
- Identifying Problems Using Data
- Data Visualization
- Presenting Data (Verbally)
- Data Driven Decisions Making (DDDM) (Making decisions based on data)

**Figure 1. Steps in Implementing Data literacy and Its Competencies**

2. Materials and Methods
The study was conducted in SMA Negeri 1 Karanganyar. The subjects in this study were 69 class XI students who were randomly selected. The data collection technique was done by giving a test to the research subject. The data collection tool is in the form of a multiple-choice type test consisting of 20 questions taken from the PISA test in 2009 [15]. This study was using a standardized test. The tests were not tied to one subject matter in the study of physics, but the test contained physics material in general. Some of the material they have gotten in-class learning is material about motion kinematics. Besides, questions related to problems in daily life and related to current issues was also included in tests such as material on global warming and renewable energy.
Data analysis techniques used qualitative analysis techniques, assisted by quantitative exposure in the form of percentage achievement of student test results in each aspects of data literacy. The aspects of data literacy was taken from the test. The results of student answers then were analyzed and categorized according to the data literacy assessment category adapted from Sudjiono, as seen in Table 1 [16].

| No. | Score  | Category      |
|-----|--------|---------------|
| 1.  | 85-100 | Very High     |
| 2.  | 70-84  | High          |
| 3.  | 55-69  | Medium        |
| 4.  | 50-54  | Low           |
| 5.  | 0-49   | Very Low      |

### 3. Result and Discussion

Data on data literacy skills of high school students in learning physics were collected using integrated tests on data literacy competencies proposed by [14]. Scoring is done based on the results of student answers to the scoring criteria that has determined. Category for scaling scale based on score 0-100 is using the calculation formula as follows:

\[
\text{score} = \frac{\text{gained}}{\text{total}} \times 100
\]

Based on the tests that have conducted at SMA Negeri 1 Karanganyar, score data was obtained for the achievement of each data literacy competency. The data was categorized according to the data literacy assessment criteria. Data on the results of the analysis in general data literacy skills of high school students, primarily in physics material, was presented in Table 2 below.

| No. | Data Literacy Competence                              | Score       | Category |
|-----|-------------------------------------------------------|-------------|----------|
| 1.  | Understanding data                                    | 63.41%      | Medium   |
| 2.  | Finding and/or obtaining data                         | 52.90%      | Low      |
| 3.  | Reading, interpreting, and evaluating data            | 65.22%      | Medium   |
| 4.  | Managing data                                         | 56.16%      | Medium   |
| 5.  | Using data                                            | 40.94%      | Very low |

Table 2. presents students’ achievement in understanding data competency as much as 63.41% with the category of "medium", finding and/or obtaining data competency as much as 52.90% with the category of "low", reading, interpreting, and evaluating data competency as much as 65.22% with the category of "medium", managing data as much as 56.16% with the category of "medium", and using data competency as much as 40.94% with the category of "very low". Based on the data obtained from the research results, data on the literacy skills of high school students were presented in graphical form in Figure 2.
In school learning, especially on physics subjects, some activities support the development of data literacy, namely experimental activities. But the reality in the field of teachers is still very rarely used to provider material and understanding concepts. Classroom learning was dominated by the teacher's role in delivering material directly. Therefore, students become familiar with the learning atmosphere. Students in calculating numbers but do not understand the meaning or interpretation of the lift. The lowest achievement of data literacy skills in competency fifth is using data with a percentage of 40.94%. It shows that students' ability to use data as evidence to support their arguments. They are accustomed to making arguments based on their own opinions or thoughts without being based on real data evidence. It is undoubtedly very influential for them later when facing the world of work that is required to be able to make a decision appropriately. Of course, it would be very fragile if the decision was only made based on opinions and ideas without being based on data evidence in the field.

The results of students' answers and arguments along with the analysis for reading, interpreting, and evaluating data competence can be seen in Table 3.

**Table 3. The Example of Student’s Answer for The Competence of Reading, Interpreting, and Evaluating Data**

| Question                                                                 | Conception               | Percentage |
|--------------------------------------------------------------------------|--------------------------|------------|
| **Correct conception**                                                   |                          |            |
| Problem 1.1: What is the average range from start to the longest straight line? | Correct conception: Distance from the starting line to the beginning of the longest straight section of the track is 1.5 km. Because on the graph shows that the longest straight section starts when car has covered a distance of 1.5 km. | 72.46%     |
| Problem 1.2: What can you say about the velocity of car racing from 2.6 km to 2.8 km? | Correct conception: The speed of the car remains constant or decreasing. | 88.41%     |
| **Incorrect conception**                                                 |                          |            |
| Problem 1.1: Student answer in addition to the 1.5 km.                    | Incorrect conception: Student answer in addition to the 1.5 km. | 27.54%     |
| Problem 1.2: The speed of the car remains constant or decreasing.         | Incorrect conception: No Response | 11.59%     |
Based on Table 3, 50 (72.46%) students answered the correct answer, and 19 (27.54%) students answered the incorrect answer in problem 1.1. and 61 (88.41%) students answered the correct answer, and 8(11.59%) students answered the incorrect answer in problem 1.2. The researchers want to know how students understand the problem to answer the question. This can be done by analyzing the arguments given by students in Figure 3a and Figure 3b.

![Figure 3](image-url)

**Figure 3.** Student Argument About Their Answer

Questions number 1, which was spread become problem 1.1 and 1.2 are examples of questions used to measure the competence of interpreting data. Student answers to both questions are correct, but students still have difficulty explaining the meaning of the graph in the question. It was seen from the reasons given by the students in problem 1.1 that do not support the answer. While the reason for problem 1.2 is enough to describe the intentions students want to express in supporting the answers, even though they still use simple language.

Through the arguments written by students after choosing the answers available, it will be known whether students really have the ability of data literacy or not. Without giving arguments in answering questions, it is possible for students to give answers inconsequentially. The results of students' answers and arguments along with the analysis for managing data and using data competence can be seen in table 4.

Based on Table 4, 32 (46.38%) students answered the correct answer, and 37 (53.62%) students answered the incorrect answer in problem 6.1. and 22 (31.88%) students answered the correct answer, 47(68.12) students answered the incorrect in problem 6.2. The researchers want to know how students understand the problem to answer the question. This can be done by analyzing the arguments given by students in Figure 4a and Figure 4b.

![Figure 4](image-url)

**Figure 4.** Student Argument About Their Answer
Problem number 6 is one example of a problem related to the current issue in everyday life. Problem 6.1 from this item is one of the questions used to measure the competency of finding data. The answers given by students indicate that students can recover data from the sources provided correctly. And problem 6.2 is one of the questions used to measure the competence of supporting argument using data. The answers given by students indicate that students' ability to argue is still lacking. Students are still easily affected by real arguments and do not check the data.

### Table 4. The Example of Student’s Answer for the Competence of Managing Data and Using Data

| Question | Conception | Percentage |
|----------|------------|------------|
| 6. THE DECREASE OF CO\(_2\) | Correct conception | 46.38% |
| From the graph above, you can see that the increasing number of percentage due to high rapid of CO\(_2\) emission is 11%. Show the calculation on how this number gotten. | Problem 6.1 | |
| Mandy analyzes the graph and claims that she found the error in percentage calculation: “The percentage of CO\(_2\) emission decrease from Europe (EU, 4%). This is not possible because Germany is a part of Europe.” | Problem 6.2 | |
| Students are able to provide arguments appropriately supported by the data presented in the diagram. | |
| Students find it difficult to find the data needed from the diagram illustrated and errors in calculating data. | |
| Students are able to give an argument, but the argument does not match the data presented in the diagram. | |
| No Response | Problem 6.1 | - |
| Problem 6.2 | - |

Based on the results of the analysis that has been carried out on the score of achievement of all data literacy competencies, the average overall percentage of data literacy skills in general is 55.72%. These results mean that students' literacy skills in general fall into medium category. There are many factors that contribute to the achievement of students data literacy competence. These factors include learning activities in schools in the form of traditional learning with the role of teachers who still dominate learning. Innovation is still very lacking in the application of learning models and the
development of learning devices used. Both students and teachers also do not know more about new competencies that are starting to be required in the digital era, one of which is data literacy.

Data literacy is very important to be invested in students since attending school. This is because data literacy is an ability that underlies the mastery of other abilities such as statistical literacy skills, critical thinking skills, analytical abilities, argumentative abilities, and others. If students succeed in developing data literacy in themselves, it will be easy for them to achieve other abilities. Having data literate students means you can use data visualizations and raw data to present material in different forms for different learners. Teaching students data literacy also helps prepare them for standardized testing and helps them understand other forms of literacy [18].

Some efforts can be made in order to develop student data literacy skills. One of them is by innovating strategies and learning methods that lead to the development of data literacy, for example learning using project-based models. Project-based learning method is one of the student-centered teaching methods which have been in use both in our country and in the world. It is one of the methods providing students with the opportunity to take part in the learning environment, making them take the responsibility of their own learning, developing students, and having them comprehend and structure information [20]. Proposed project usually have problems related to practical fact about the content of the proposed course [21].

Erwin in his research explained that by using real-world data sets in project based instructional unit, teachers can guide students in analyzing, interpreting, and reporting quantitative data. Working with authentic unfiltered, and previously wasn’t interpreted data helps students to integrate data literacy with their disciplinary learning.[17]

The learning center is also more emphasized on the role of students, for example with experimental activities will try directly and kinetic with data. Through discussion activities from the results of experiments conducted also train students in arguing. However, students must still be guided by the teacher so that the argument presented contain the basis of the data they found.

4. Conclusion and Suggestion
Based on the research that has been done, it can be concluded that the lowest percentage of data literacy is in the competence of using data. The acquisition of score was 40.94% with the category of "very low". The highest percentage of data literacy is in the competence of reading, interpreting, and evaluating data with the acquisition of score was 65.22% in the category of "medium". The average data literacy score of high school students in physics learning in general for overall competency is 55.72%. The average score is included in the "medium" category.

Senior high school students should have good data literacy skills. This is because data literacy is one of the competencies that are starting to be required when students enter the workforce as well as taking higher education. One effort that can be done to improve data literacy skills is through improving the learning process in schools such as the development of learning devices by applying various learning models that are integrated with data literacy skills so that students are able to join the global challenge in 4.0 Industry.

Acknowledgements
The authors would like to express their thanks to the schools and educators that participated in the research. The authors would also like to thank the educators who provided assistance during the research. Furthermore the authors give special thanks to our parents and my husband that motivated and prayed for us.

References
[1] Tibor Koltay 2014 Big Data (Big Literacies? Tema. Vol 24) p 3-8
[2] Harris J 2012 Data Is Useless without The Skills to Analyze It (Harvard Business Review)
[3] Ridsdale Chantel et al 2015 Strategies and Best Practices for Data Literacy Education (Kanada: Dalhousie University)
[4] Jordan and Morrow 2018 Developing A Data literate Workforce: A Strategy and Framework for the Enterprise (Data Literacy Program QlikTeach International AB)

[5] Australian Public Service Commission 2016 Core Skill Data Literacy Guide (Australia)

[6] Agodfrey 2012 Data Literacy (Public Schools of North Carolina, State Board of Education Department of Public Instruction)

[7] Education Development Center 2016 Building Global Interest in Data Literacy: A Dialogue (Workshop Report) EDC (Ocean of Data Institute United States of America)

[8] Gunter G A 2017 Building Student Data Literacy: An Essential Critical-Thinking Skill for The 21st Century MultiMedia & Internet@Schools (H.W.Wilson-EDUC vol 14(3)) p 24-28

[9] Qin J and D’Ignazio J 2010 Lessons Learned from a Two Year Experience in Science Data Literacy Education (International Association of Scientific and Technological University Librarians, 31st Annual Conference)

[10] Wing J 2008 Computational Thinking and Thinking about Computing (Philosophical Transactions of The Royal Society Vol 366) p 3717-3725

[11] Davenport T and Patil D J 2012 Data Scientist: The Sexiest Job of the 21st Century (Harvard Business Review)

[12] Deahl Erica 2007 Better the Data You Know: Developing Youth Data Literacy in Schools and Informal Learning Environments (Departement of Comparative Media Studies, Yale University)

[13] Carlson J et al 2011 Determining Data Information Literacy Needs: A Study of Students and Research Faculty (Portal: Libraries & the Academy vol 11(2)) p 25–26

[14] Prado J C and Marzal M Á 2013 Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents (De Gruyter. Syracuse University Library)

[15] OECD 2009 Take the Test Sample Question from OECD’s PISA Assesments (http://oecd.org/publishing/corrigenda)

[16] Sudijono A 2009 Pengantar Evaluasi Pendidikan (Raja Grafindo Persada, Jakarta)

[17] Erwin R W 2015 Data Literacy: Real-World Learning Through Problem-Solving With Data Sets (American Secondary Education 43(2))

[18] Jennifer C 2017 Where The Rubber Meets The Road: Data Litaercy in The Content Areas (Data Literacy in The Real World: Conversation and Case Studies)

[19] W B Kippers et al 2018 Data Literacy: What Do Educators Learn and Struggle with During A Data Use Intervention? (Studies in Educational Evaluation,Vol. 56) p 21-31

[20] N R Ergul and E K Kargin 2014 The Effect Of Project Based Learning On Students’ Science Success (Social and Behavioral Sciences Vol. 136) p 537 – 541

[21] S Amamou and L C Belcadhi 2018 Tutoring in Project-Based Learning (Procidea Computer Science Vo. 126) p 176-185