

# Towards categorizing ethical questions in data literacy


## Results of a focus groups study at the NFDI4Ing conference 2022

Samira Khodaei  <sup>1</sup>Anas Abdelrazeq  <sup>1</sup>Ingrid Isenhardt  <sup>1</sup>

1. Chair of Production Metrology and Quality Management Information Management in Mechanical Engineering organization,

**Date Received:**

2022-12-16

**Licenses:**This article is licensed under: **Keywords:**

ethical literacy, data literacy, ethics, interdisciplinary research, focus group study

**Data availability:**

Data was not created through the work on this paper

**Software availability:**

Software was not created through the work on this paper

**Abstract.** Data Literacy is crucial for a sustainable engineering education [1]. In aiming to find solutions to solve future challenges, mechanical engineering has started to integrate data literacy into the higher education curriculum [2]. However, ethics are rarely considered in current frameworks. Ethics are seen as a side topic or are equated to data privacy issues [3]. Since literacy aims to empower people to make informed decisions based on their or other data[4], the development of critical reflection and discussion on ethics is central for data literacy. In our contribution, we will first summarize current existing data literacy frameworks and their ethics concept. Then, through a focus group study among data literacy experts' ethical questions in data literacy are collected and categorized. The study was conducted with 15 experts at the NFDI4Ing Conference 2022. This approach expands ethical issues in data literacy beyond data privacy towards applied, current and pressing ethical topics.

## 1 Introduction

2 Describing a 'set of abilities around the use of data as part of everyday thinking and reasoning for  
3 solving real-world problems'[5], data literacy is key for an increasingly data driven society [6].  
4 Along with the ability to solve real-world problems with the use of data, the critical reflection  
5 with data is becoming increasingly important [7]. Moreover, there are many risks of incorrect  
6 assumptions based on data that might lead to incorrect knowledge and decisions. This then might  
7 further fuels biases in societies. There is a responsibility for those communicating through data  
8 to inadvertently reduces biases [5].

9 Ethics systematize, defend, and recommend concepts of right and wrong behavior and action  
10 [8]. This often results in extensive discussions of complex, interdisciplinary and ambiguous  
11 questions – especially in an increasing dynamic and complex global society. To become agent  
12 in their decision making, ethical guidelines based on democratic values had been introduced  
13 in other literacies such as media literacy [9]. As the relevance of data increased along with the  
14 difficulty for human beings to comprehend the influence on our knowledge and decision-making,  
15 ethics need to be further considered in the data literacy frameworks.

16 Ethical considerations should not be understood as a side subject to be taught with many others,

17 but as a core element and basis for all subsequent decision making. Especially competencies  
18 that consider critical thinking and enabling agency are barely mentioned in current literacy  
19 frameworks. As the relevance of data increased along with the difficulty for human beings to  
20 comprehend and process, the influence to our knowledge culture should be further considered in  
21 the frameworks.

22 Indeed, many literacy discussions consider ethical discussions as important for supporting  
23 empowered citizens [2] [9] [10]. Still, when applying ethics in the curricular topics of data  
24 literacy, they are often pushed to the side in favor of more applicable topics such as data  
25 visualization, data analytic or data tasting.

26 The aim of this contribution is to spotlight ethics in data literacy frameworks. The research  
27 question is therefore:

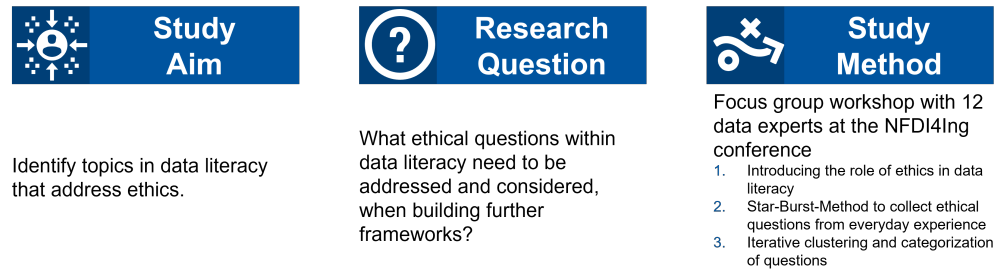
- 28 • What ethical questions within data literacy need to be addressed and considered, when  
29 building further frameworks?

30 The first part of this contribution will highlight ethics concepts in existing data literacy frame-  
31 works. Through, this literature overview a concept how to position ethics in data literacy models  
32 will be introduced. In order to become more concrete, the following part will introduce a focus  
33 group study as a tool to collect pressing ethical topics in the interaction with data. The focus  
34 group study took place among data literacy experts at the NFDI4Ing conference in November  
35 2022. The different ethical questions are summarized to identify key ethical categories that  
36 need to be included in data literacy. Finally, the conclusion will open further potential research  
37 questions in data literacy.

## 38 2 The role of ethics in data literacy frameworks

39 Contrary to its importance in decision making, ethics remain a minor course within data literacy.  
40 They rarely play the central role that is required. Most of the current frameworks that do consider  
41 data ethics as important then lack concrete applicable topics in their curricula. They rarely  
42 are concrete and give hints to educators on how exactly they can apply ethics in data literacy  
43 programs.

44 For example, Heidrich et. al. introduce ethics as a side competency in their framework [11].  
45 In the study from Wolff et al, they identify through card sorting that professionals see ethical  
46 competence as highly relevant within data literacy, but do not give further examples on what  
47 asked professionals understand by this [5]. In Grillenberges and Romeikes approach to create a  
48 data literacy Competency Model based of Risdale et al, they introduce their competencies along  
49 the data management cycle and divide them into process and content-oriented competencies  
50 [12]. They introduce a layer called ethics, but do not connect it visibly with the introduced  
51 competencies or exemplify it. Schüller et al introduce a comprehensive data literacy framework  
52 considering both comprehensive and selective competencies along a data value chain [2]. In  
53 their model ethics is pushed to the side of the framework and is seen as a separate ethics literacy.  
54 Closest to concrete examples in ethics is the research team around Giese. They introduce ethics  
55 as part of the transparency and awareness pillar [10]. This pillar is one of three other pillars and  
56 additionally includes a law and technical component. The ethical pillar in the concept of Giese

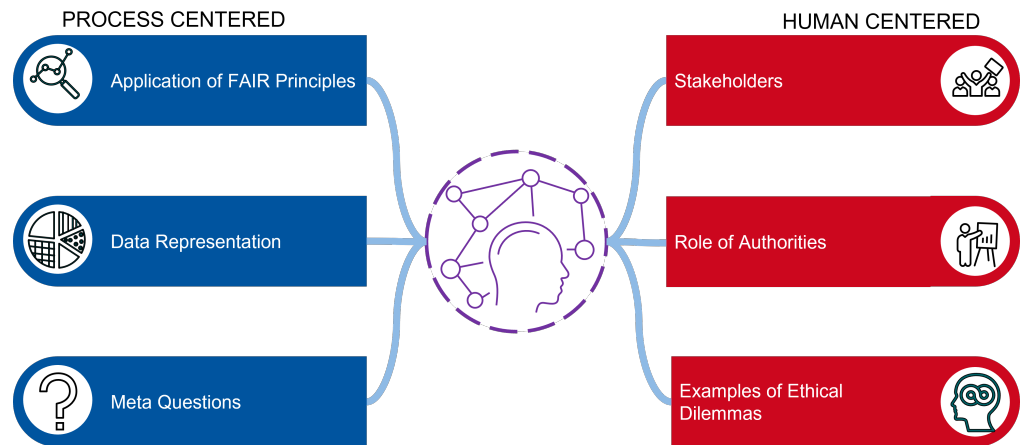


**Figure 1:** Overview on focus group study design

57 et al was introduced through real-world examples and thinking and pairing exercises [10].  
 58 In their example, they introduce a case from twitter, which indicates the importance of (social)  
 59 media understanding, when it comes to ethics in data literacy. The example of Giese's application  
 60 of ethics reveals that ethics within data literacy is often connected to other literacy types. This  
 61 might be the reason why Schüller et. al. frameworks introduce ethics as an additional literacy in  
 62 their concept.  
 63 Ethical considerations in data literacy should be seen as a core element for all subsequent decision  
 64 making. They should not merely be applied at some point in the process, but always remain in the  
 65 core of a data literacy concept. Regardless of the data processing step aside from the *how* there  
 66 should always also be the question of the *why*. As ethical questions require the consideration of  
 67 a wide range of stakeholders and other fields, therefore ethical questions are usually overlapping  
 68 with other literacy concepts.

### 69 3 Method focus group study

70 To answer the research question concerning the content of the ethic topics, a focus group study  
 71 with data literacy experts and professionals was conducted. Focus group studies are a qualitative  
 72 discourse method in which a group is stimulated to discuss a specific topic [13]. While the  
 73 researchers provide a specific focus, such as ethics in data processing, the data is collected  
 74 through the observation of a groups response through this topics. According to Kitzinger, this  
 75 method is used to generate and explore questions among a group and encourage the development  
 76 of their own analysis of common experiences [14]. While this method might not give a deep  
 77 insight into individual perspectives and experiences[13], it is well suited to identify norms and  
 78 values based on a common experience within a group [14].  
 79 Therefore, this method has been selected to gain a further understanding of ethical dilemmas  
 80 among a group of data experts (see figure 1). Due to being a complex topic, ethical dilemma  
 81 are a helpful to identify shared experiences in the decision making process of data. Through the  
 82 discussion in groups, the individuals might find solutions or at least see that there are patterns to  
 83 their experienced dilemma. This is helpful for developing a collection of applied ethic topics  
 84 that go beyond the usual questions of data privacy.  
 85 The focus group study was conducted at the NFDI4Ing conference in October 2022 to a group of  
 86 15 participants with various background in mechanical engineering, information science and  
 87 software engineering. After an exchange about the urgency of ethics and agency, the starburst  
 88 method was introduced to collect ethical questions from the experts in smaller rotating groups.



**Figure 2:** Overview on categorized results of focus groups study

89 The star bursting method is a method in design thinking to collect questions in order to understand  
 90 a problem from different perspectives [15]. In this method a star with six spikes represents six  
 91 question words (how, who, why, what, when and where). The task for the participants is to  
 92 reflect and fill the question words with ethical questions they have faced in their professional  
 93 work with data.

94 The group was divided into two groups and asked to collect and discuss ethical question based  
 95 on the six question words. The idea behind the ethical questions was not focused on finding  
 96 solutions at this point, as it is the nature of such questions to not be easily answerable from the  
 97 point of one domain. Rather, this collection was useful in understanding the spectrum of ethical  
 98 questions and the contexts that need to be considered when working with data. These questions  
 99 were subsequently anonymized and categorized and are presented in the following part.

### 100 3.1 Results of the focus group study - categorizing data ethics

101 Through this study around 20 ethical questions in data focused research were collected among the  
 102 experts. While the explicit answering of these questions was not the aim of the study, the different  
 103 considerations help to gain an understanding of ethical aspects that need to be considered when  
 104 addressing ethical questions in data literacy.

105 The ethical questions were summarized in the following six categories (See figure 2): the appli-  
 106 cation of the FAIR Principles (4), Stakeholders (4), Role of Authorities (3), Data Representation  
 107 (3), Ethical Dilemmas and Examples (3), and a category consisting of questions that did not fit  
 108 the other categories (2).

109 The FAIR principles are findability, accessibility, interoperability, and reusability (FAIR) [16].  
 110 An example question based on these principles was *'When should data transparency be given  
 111 and when is it too much?'*. As transparency is an underlying theme and the central aim of the  
 112 FAIR principles, this question addresses an important decision that people working with data are  
 113 considered daily.

114 The Stakeholder category reflects different groups that are affected by data-based applications.  
 115 The question pair reflecting this is *'Who might struggle with such ethical standards?'* and *'Who  
 116 would mainly benefit from such ethical standards?'*. This category has an overlap with both the

117 role of authorities and data representation categories.  
118 The role of authorities has evolved around the power that states and companies hold. An example  
119 question here was, *'Where can I turn to with an ethical dilemma in data?'* in combination with  
120 *'Who could have the responsibility for deploying ethical standards in different application areas*  
121 *(e.g. research, practice)?'*. This is more of a meta-category describing the organization of ethics  
122 rather than their application, which is reflected more in the Stakeholder category. It might be a  
123 subcategory of the Stakeholder category but is presented here as a separate category due to the  
124 amount of questions that arise in the discussion.  
125 Data Representation overlaps with Stakeholders and includes questions like *'What can we do*  
126 *against misinterpretation of data?'* and *'How can we show that data representation reflects the*  
127 *truth?'*. This category is strongly connected to practical guidelines in design and visualization.  
128 As the visualization of data is closely connected to visual and media literacy, those ideas might  
129 be found in overlapping areas of the other literacies.  
130 The Ethical Dilemmas and Examples category collected questions from concrete, applied exam-  
131 ples in daily life. An example question for the category is *'How can we detect bias in data?'*.  
132 The further collection of examples would be helpful for a concrete design of an educational  
133 curriculum, as this category tends to become more specific than the others. There were further  
134 ethical questions that were sorted into the remaining collected category, such as *'When should*  
135 *data literacy and ethical maturity be taught?'*, which is more oriented towards education, and  
136 *'How could Ethics impede data content generation?'* as further practical ethics questions. As  
137 this is a first attempt to address the variety of ethical questions in data management, further focus  
138 studies might develop further categories based on those questions.  
139 Finally, in a reflection and feedback round of the study, the exchange gave new insights for  
140 the group as well as for the data. The biggest downside addressed by the group was that this  
141 exchange was too short and could have been extended further. Still, the collected categories  
142 extend current ethics in data literacy with a collection of topics that professionals recently face.

#### 143 **4 Conclusion and outlook**

144 This paper aims to broaden the understanding of data literacy by including discussions and  
145 critiques from media literacy into the development of a data literacy framework. This approach,  
146 with a literature review in combination with a focus group study among data literacy experts,  
147 can be seen as a step towards developing agency and ethical foundations in literacy frameworks  
148 that go beyond data privacy discussions. To understand the research questions

- 149 • What ethical questions within data literacy need to be addressed and considered, when  
150 building further frameworks?

151 First a literature study compared how different data literacy concepts applied ethics in their  
152 frameworks. As a result, it became clear that ethics is often seen as important but is rarely  
153 prominently applied. Additionally, it was concluded that applied reflection of ethical questions  
154 need to include multiple perspectives. Still, the shift of ethics into the center is required, as  
155 ethical considerations are not limited to one scientific field.

156 To fill ethics in data literacy, a focus group study was conducted among data literacy experts

157 at the NFDI4Ing conference in October 2022. Through an online workshop around 20 ethical  
 158 questions were collected, categorized, and introduces (see figure 1). The main categories are the  
 159 Application of FAIR Principles, Stakeholders, Role of Authorities, Data Representation, Ethical  
 160 Dilemmas and Examples (see figure 2). These questions give further insights into themes that  
 161 ethical programs in Data literacy apply and which are worth further examination.

162 As a next step, the scientific exchange between different literacy framework is highly recom-  
 163 mended. Some of the collected ethical questions overlap other scientific fields such as media or  
 164 sustainability literacy. Through further interdisciplinary exchange, data literacy will empower  
 165 professionals, students and educators to make informed data-based decisions.

## 166 5 Acknowledgements

167 This paper would not be possible without opportunity to conduct a workshop at the NIDI4Ing  
 168 conference 2022. The authors thank the participants of the workshop for their active and engaged  
 169 contribution to the ethical questions, insights in the ethical dilemmas they have faced in their  
 170 professional life and the open exchange on eye-level.

## 171 6 Roles and contributions

172 **Samira Khodaei:** Conceptualization, Execution, Writing, Original Draft

173 **Anas Abdelrazeq:** Review & Editing

174 **Ingrid Isenhardt:** Review & Editing

## 175 References

- 176 [1] M. Richter, “Fair data management in engineering sciences in the first semester of the  
 177 bachelor of mechanical,” in *Data literacy right from the start*, R. Schmidt, Ed., 2022.
- 178 [2] K. Schüller, “Ein framework für data literacy,” *AStA Wirtschafts- und Sozialstatistisches*  
 179 *Archiv*, vol. 13, no. 3-4, pp. 297–317, 2019, ISSN: 1863-8163. DOI: [10.1007/s11943-](https://doi.org/10.1007/s11943-019-00261-9)  
 180 [019-00261-9](https://doi.org/10.1007/s11943-019-00261-9).
- 181 [3] T. Braun, A. Büsch, V. Dander, *et al.*, “Positionspapier zur weiterentwicklung der kmk-  
 182 strategie <bildung in der digitalen welt>,” *MedienPädagogik: Zeitschrift für Theorie und*  
 183 *Praxis der Medienbildung*, pp. 1–7, 2021. DOI: [10.21240/mpaed/00/2021.11.29.X](https://doi.org/10.21240/mpaed/00/2021.11.29.X).
- 184 [4] M. Rowe, Ed., *The journal of media literacy*. Madison: National Telemedia Council Inc.,  
 185 2017, vol. 64.
- 186 [5] A. Wolff, D. Gooch, J. J. Cavero Montaner, U. Rashid, and G. Kortuem, “Creating  
 187 an understanding of data literacy for a data-driven society,” *The Journal of Community*  
 188 *Informatics*, vol. 12, no. 3, 2016, ISSN: 1712-4441. DOI: [10.15353/joci.v12i3.3275](https://doi.org/10.15353/joci.v12i3.3275).
- 189 [6] D. Taibi, L. Fernandez-Sanz, V. Pospelova, *et al.*, “Developing data literacy competences  
 190 at university: The experience of the dedalus project,” in *2021 1st Conference on Online*  
 191 *Teaching for Mobile Education (OT4ME)*, IEEE, 2021, pp. 112–113, ISBN: 978-1-6654-  
 192 2814-9. DOI: [10.1109/OT4ME53559.2021.9638912](https://doi.org/10.1109/OT4ME53559.2021.9638912).

- 193 [7] B. Motyl, G. Baronio, S. Uberti, D. Speranza, and S. Filippi, “How will change the  
194 future engineers’ skills in the industry 4.0 framework? a questionnaire survey,” *Procedia*  
195 *Manufacturing*, vol. 11, pp. 1501–1509, 2017, ISSN: 23519789. DOI: [10.1016/j.prom](https://doi.org/10.1016/j.promfg.2017.07.282)  
196 [fg.2017.07.282](https://doi.org/10.1016/j.promfg.2017.07.282).
- 197 [8] J. Ferretti, Daedlow K., J. Kopfmüller, M. Winkelmann, A. Podhora, and Walz, R., Bertling,  
198 *Reflexionsrahmen für forschen in gesellschaftlicher verantwortung*, Berlin, 2016.
- 199 [9] M. Leaning, *Media and information literacy: An integrated approach for the 21st century*  
200 (Chandos information professional series). Cambridge, MA and Kidlington: Chandos  
201 Publishing an imprint of Elsevier, 2017, ISBN: 9780081002353. [Online]. Available:  
202 <https://aml.ca/wp-content/uploads/2017/03/JMLVo.64No.12-2017.pdf>.
- 203 [10] T. G. Giese, M. Wende, S. Bulut, and R. Anderl, “Introduction of data literacy in the  
204 undergraduate engineering curriculum,” in *2020 IEEE Global Engineering Education*  
205 *Conference (EDUCON)*, IEEE, 2020. DOI: [10.1109/educon45650.2020.9125212](https://doi.org/10.1109/educon45650.2020.9125212).
- 206 [11] J. Heidrich, P. Bauer, and D. Krupka, “Future skills: Ansätze zur vermittlung von data  
207 literacy in der hochschulbildung,” *Hochschulforum Digitalisierung*, no. 37, 2018.
- 208 [12] A. Grillenberger and R. Romeike, “Developing a theoretically founded data literacy  
209 competency model,” in *Proceedings of the 13th Workshop in Primary and Secondary*  
210 *Computing Education*, New York, NY, USA: ACM, 2018. DOI: [10.1145/3265757.326](https://doi.org/10.1145/3265757.3265766)  
211 [5766](https://doi.org/10.1145/3265757.3265766).
- 212 [13] M. Schulz, “Quick and easy!? fokusgruppen in der angewandten sozialwissenschaft,” in  
213 *Fokusgruppen in der empirischen Sozialwissenschaft*, M. Schulz, B. Mack, and O. Renn,  
214 Eds., Wiesbaden: Springer VS, 2012, pp. 9–22, ISBN: 978-3-531-19396-0.
- 215 [14] Jenny Kitzinger, “Qualitative research: Introducing focus groups,” *BMJ*, vol. 311, no. 7000,  
216 pp. 299–302, 1995, ISSN: 1468-5833. DOI: [10.1136/bmj.311.7000.299](https://doi.org/10.1136/bmj.311.7000.299). [Online].  
217 Available: <https://www.bmj.com/content/311/7000/299>.
- 218 [15] S. Blakely, *Starbursting technique: How to brainstorm using starbursting*, masterclass.com,  
219 Ed., 2022. [Online]. Available: [https://www.masterclass.com/articles/starbur](https://www.masterclass.com/articles/starbursting)  
220 [sting](https://www.masterclass.com/articles/starbursting) (visited on 12/07/2022).
- 221 [16] European Commission, *A european strategy for data: Com(2020) 66 final*, 2020. [Online].  
222 Available: [https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri](https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0066&from=EN)  
223 [=CELEX:52020DC0066&from=EN](https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0066&from=EN) (visited on 12/07/2022).