Extended Abstract for the
Symposium on Integrating AI and Data Science into School Education Across Disciplines 2025

FOSTERING DATA AND AI COMPETENCIES IN PRIMARY SCHOOLS

Anja Gärtig-Daugs[1] & Ute Schmid[2]
[1]University of Bamberg, Germany, anja.gaertig-daugs@uni-bamberg.de
[2]University of Bamberg, Germany

Focus Topics: AI and data science curricula and implementation in schools, learning materials

Relevance of data and AI competencies in primary schools

Data-driven tools and applications of Artificial Intelligence (AI) are shaping our everyday lives and are therefore present in children's daily lives: Almost 50% of 6 to 13 year olds use their mobile phones every day to browse the internet, share on social networks, and stream videos or music. One third use intelligent voice assistance systems (Feierabend et al., 2023). In Germany, a key goal of education policy is to support children in exploring, understanding and shaping the digital world through appropriate educational programs (KMK 2016; KMK 2021). In this context, it is important to promote an understanding of basic principles of data science and AI (Schmid et al., 2021). Children should know how AI systems function and that data build a crucial basis for the training of AI technologies. They should be aware that the data they reveal about themselves by interacting with search engines, intelligent voice assistants or large language models can be used to make inferences about the individual users. They should also be aware that data can be used to improve AI systems or, in the worst case, to manipulate users. Therefore, we argue that the conceptual basics of data literacy and AI should be integrated in primary school lessons.

Previous research has shown that key concepts of data literacy can be didactically reduced and conveyed in a way that is suitable for primary school children. For example, Wolff et al. (2016) taught children aged 9 to 10 years how to analyse, interpret and reflect on data from smart electricity meters. A prerequisite for primary school children to acquire forerunner skills on their way from mere users to responsible questioners of data and AI is that teachers provide appropriate learning opportunities. Despite the recommendation of the German Informatics Society (2023), data and AI competencies are not yet a compulsory part of (university) teacher training. As a result, many teachers lack data and AI literacy (Lindner & Berges, 2023; Vo & Pancratz, 2023), which would be a prerequisite for methodological and didactic approaches to foster data and AI competencies among students.

Professionalising primary school teachers to promote Data and AI literacy through a Massive Open Online Course (MOOC)

Theoretical conception of the course

Data literacy encompasses the areas of knowing and understanding data and data sources, collecting data, managing data, evaluating data, applying data, and critically reflecting on the handling of data (Ridsdale et al., 2015; Grillenberger & Romeike, 2018). Data skills are fundamental for the understanding of AI systems and their underlying databases (Schmid, 2025). They are therefore are an integral part of AI literacy. AI literacy refers to the ability to know and understand, use and reflect on AI concepts, technologies, applications, and their implications for individuals, society and the environment (Ng et al, 2021; Schmid, 2025). Relevant topics for AI education in primary schools can be derived from the Five Big Ideas of AI (Touretzki, Gardener-McCune & Seehorn, 2019): perception, representation and reasoning, machine learning, natural interaction, and reflecting on the positive and negative societal impacts of AI technologies. Teachers can introduce these ideas from primary school onwards. This spiral curriculum approach allows for initial experiences and explanatory approaches, and enables students to experience the complexity of systems step by step (Diethelm, 2021).

From a didactic point of view, computer science education in primary schools should relate to children's concrete experiences (GI, 2018). Learning situations should be designed in a way that allows children to learn about the functioning principles and the underlying theoretical concepts. Furthermore, they should provide concrete user experiences with analogue playing materials or digital applications. Finally, they should encourage reflection on the implications for the individual and for society as a whole (Brinda et al., 2016; GI, 2018; Diethelm, 2021; Schmid et al., 2021)

In S. Podworny & S. Schönbrodt (Eds.), *Towards Fostering AI and Data Science Literacy in Schools Across Disciplines. Proceedings of the 1st Symposium on Integrating AI and Data Science into School Education Across Disciplines (AIDEA1 2025)*, Salzburg, Austria.

Based on these theoretical considerations, we developed a massive open online-course to qualify primary school teachers for integrating data and AI topics in primary school lessons. The course can either be integrated into university course in a blended learning format, or can be taken as a self-study course by professional teachers. According to the findings of Lindner & Berges (2023) and Vo & Pancratz (2023), no prior knowledge is required to follow the course. With regard to the relevant teacher competencies for teaching in a digitalized world (Schultz-Pernice et al., 2017), the course aims at fostering technical and didactic competencies in the field of data science and AI, and empowers teachers to plan, conduct, evaluate and share lessons that deal with data and AI topics. As a massive open online course, the course is also open to interested third parties.

Structure of the course

Following to the data literacy framework mentioned above, the course is structured in five modules. From the Five Big Ideas of AI, the course covers the areas of machine learning and natural interaction and engages the reflection on the implications of these technologies.

Each module starts with an introduction to the theoretical background for teachers. To support self-directed learning, the course includes explanatory short videos, information texts, interactive presentations and exercises, as well as self-assessment quizzes based on H5P (HTML5 Package; H5P Group, 2024).

At the end of each module, didactic suggestions and game activities are provided to facilitate the integration of the topics into the classroom. The materials are designed to promote data and AI literacy in different contexts and subjects such as Maths, Language, Science or Arts. This illustrates the interdisciplinary relevance and potential of data and AI competencies. All materials are CC-By-licensed and can be adapted by teachers to suit the interests and needs of their students.

Overview about the modules and the teaching materials

The first module “Data, Information and Knowledge” provides the conceptual framework for the following modules. It clarifies basic terms such as data, information, and knowledge. It deals with different types of data, the analysis and visualization of data. To illustrate the theoretical concepts and to enable teachers to use the examples in the classroom, the course uses fictional data from the children`s world, such as fact sheets on students, as well as real world data from official statistics, for example on pets. The didactic suggestion includes teaching materials that illustrate the difference between data and information by using fictional fact sheets of a fantasy character. Inspired by this example, the children are asked to fill in their own fact sheet about themselves and thus collect data.

The second module “Data Storage and Data Access” introduces the concept of digital representation of different types of data such as numeric, text and image data, and deals with the advantages and disadvantages of digital data storage. The teaching materials includes worksheets that illustrate how images can be represented using pixels, which are stored in the computer as a sequence of zeros and ones. This familiarizes children with notation using zeros and ones as information carriers, which can later be transferred to other types of data.

The third module “Learning from Data” deals with the possibilities and limitations of learning from data using machine learning methods. It discusses the particularities of learning algorithms that are trained on large data sets. As exemplary machine learning methods, the perceptron for the binary classification of objects and the functioning of a chatbot that can simulate human conversation are illustrated. The didactic inspiration includes a game that allows children to experience how computers can be trained by means of given examples to make predictions about new objects from given data (Schmid et al., 2021). The objective of the game is to identify relevant features of a package, such as size or wrapping paper, that indicate the presence of a coin in the packet. For this purpose, certain parcels may be checked. The materials are designed for teams of two: Student 1 closes his or her eyes. Student 2 hides the coins in all the packages with dots as indicated in the first row of the template (see Figure 1). When all the coins are hidden, Student 2 pushes up the packets marked with an arrow, as indicated in the second line of the template. Now, Student 1 can now open his or her eyes and check whether the packages that have been pushed up contain a coin.. By comparing the characteristics of the packages that contain a coin, he must decide which other packages contain a coin. In the given example, only the packages with dotted wrapping paper contain a coin. The size of the packages is irrelevant.

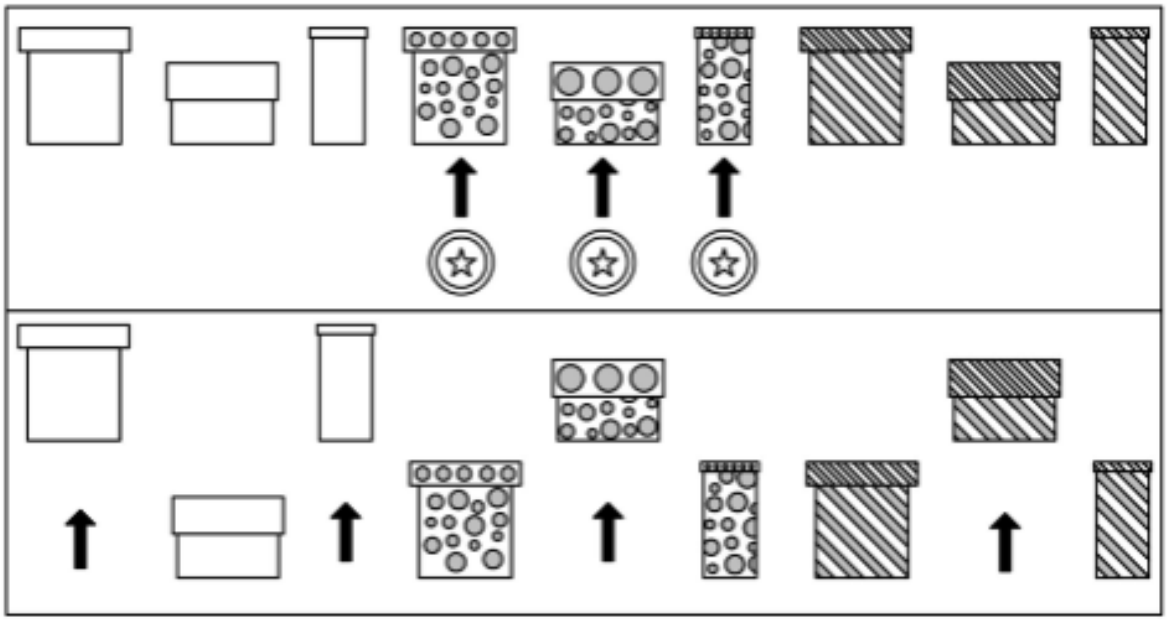

Figure 1: Template for preparing the coin search activity (AI campus course "Data literacy for Primary Schools", Module 3, https://ki-campus.org/courses/dlgrundschule)

The fourth module "Reliability of Data" looks at different types of error and bias, discusses the social problems associated with biased data, and provides strategies for checking the plausibility of information. The teaching material introduce the use of measuring instruments such as a thermometer. By writing down room temperatures they measure, children can observe that measurement data is subject to random bias and can fluctuate. They can also made be aware that data can be affected by deliberate bias, for example, when the thermometer is manipulated by breathing on it or using a hair dryer.

Finally, the fifth module "Privacy and Data Security" deals with the difference between privacy and data security, raises awareness of the (economic) interests behind data collection, and introduces the basic principles of strong passwords and data encryption. The didactic suggestion appeals children's combinatorial skills and illustrates that the longer a password is and the more different characters are available to create it, the more secure it is.

Empirical evaluation of the course

The course is accompanied by a short questionnaire that aims to provide insights into the effects of the course on the development of data and AI-related knowledge and the perceived data and AI-related teaching skills after following the course. Therefore, all participants are asked to participate in a short survey at the beginning and at the end of the course.

From the start of the course in March 2022 to September 2024, 340 participants had enrolled in the course. All responded to the introductory questionnaire, but only 89 took part in the end-of-course survey. The completion of both questionnaires is voluntary. Only the results for people who took part in both surveys are presented below (53 female, 26 male, 1 diverse, 9 refused to answer; mean age = 31.4 years, range: 9 to 72 years; 61.8% are pre-service or professional teachers).

At the beginning of the course, the participants are given a brief overview of the content they can expect, i.e. the five modules and the learning objectives are listed. Participants are then asked to assess their prior data and AI-related knowledge measured by the question "How would you rate your prior knowledge of the course topics?". They can answer this question on a 5-point Likert scale (1: very extensive - 5: very limited). At the end of the course, the same question is asked again to measure the knowledge of data and AI after the course. Figure 2 shows the results of the pre- and post-assessment: Most participants (34.8% out of 89 persons) rated their prior knowledge at a medium level (mean =

3.09, standard deviation = 0.97). After the course, most participants (60% out of 89 persons) rated their knowledge of data and AI as rather extensive (mean = 1.9, standard deviation = 0.57).

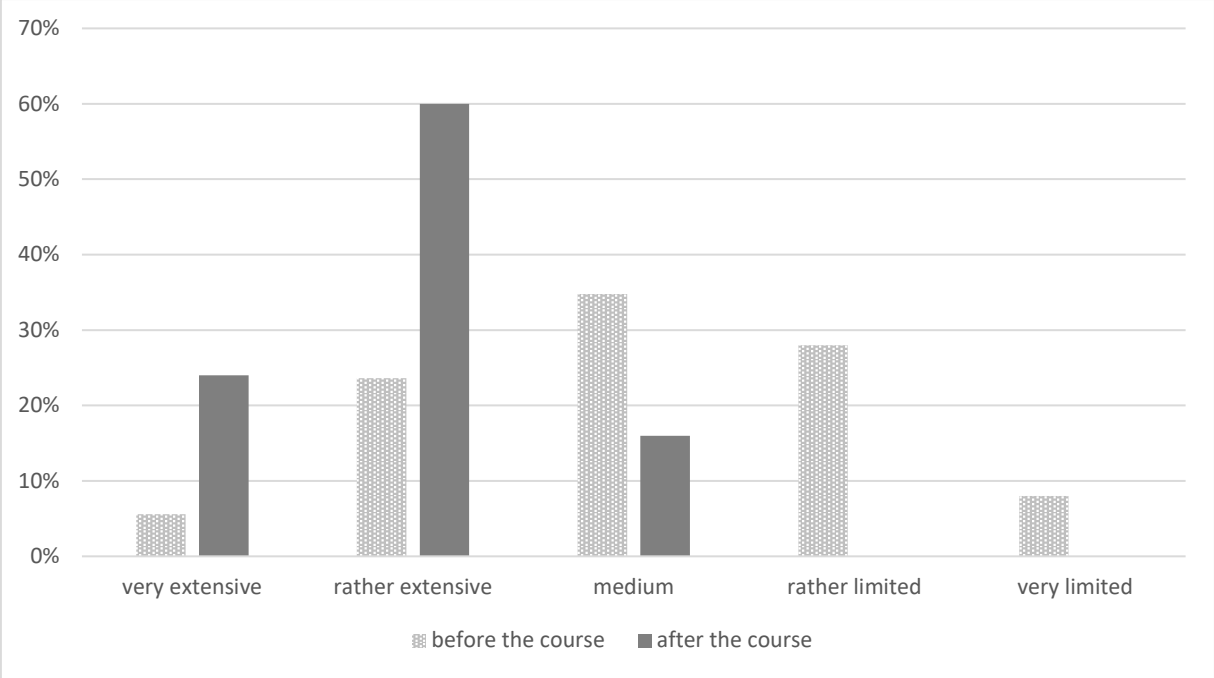

Figure 2: Self-assessed knowledge of data and AI knowledge before and after the course

At the end of the course, the participants were also asked to rate their perceived self-efficacy to explain the course topics and thus to teach data literacy skills. This was measured by the question "Do you dare to explain the course topics to others?" Participants were again able to rate their answers on a 5-point Likert scale (1: completely - 5: not at all). 76% (52 out of 68 persons who answered this question) said that they would be able to explain the topics completely or at least to a large extent (mean = 1.99, standard deviation = 0.70). No one said that they would not be able to explain it to others.

The results suggest that a massive open online course may be a suitable way to promote data and AI-knowledge among pre-service and professional teachers as well as other interested persons. After finishing the course, the huge part of the participants is convinced to be able to explain the topics covered by the course to others. Thus, a massive open online may be a good way to prepare teachers for the integration of data and AI literacy in primary school education.

Further research should also focus on the perceived relevance of data and AI literacy in primary education and the planned behavior of primary school teachers to integrate these topics into their teaching. In addition to the perceived self-efficacy to teach data and AI in the classroom, an instrument to objectively measure teachers' competencies before and after attending the course would provide deeper insights into the training effects of the course.

References

Brinda, T., Diethelm, I., Gemulla, R., Romeike, R., Schöning, J., & Schulte, C. (2016, December 15). *Dagstuhl-Erklärung: Bildung in der digital vernetzten Welt.* https://dagstuhl.gi.de/fileadmin/GI/Hauptseite/Aktuelles/Projekte/Dagstuhl/Dagstuhl-Erklaerung_2016-03-23.pdf

Diethelm, I. (2021, June 15). *Künstliche Intelligenz und Informatik.* https://digid.jff.de/ki-expertisen/ki-informatik/

Feierabend, S., Rathgeb, T., Kheredmand, H., & Glöckler, S. (2023). *KIM-Studie 2022. Kindheit, Internet, Medien. Basisuntersuchung zum Medienumgang 6- bis 13-Jähriger*. Medienpädagogischer Forschungsverbund Südwest.

GI/Gesellschaft für Informatik (2018, July 9). *Kompetenzen für informatische Bildung im Primarbereich.* http://ddi.uni-wuppertal.de/website/repoLinks/v63_Kompetenzen_Entwurfsfassung_2018-06-26.pdf

GI/Gesellschaft für Informatik (2023, November 29). *Informatikkompetenzen für alle Lehrkräfte*. https://dl.gi.de/server/api/core/bitstreams/e271f87d-1087-4617-83f5-b0ef98984b04/content

Grillenberger, A., & Romeike, R. (2018). Data Literacy und das Modell der Schüsselkonzepte des Datenmanagements. In Gesellschaft für Informatik (Ed.), *Data Literacy und Data Science Education: Digitale Kompetenzen in der Hochschulausbildung* (pp. 10–13).

H5P Group (2024). *H5P* [HTML5 content]. https://h5p.org/

KMK/Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2016). *Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz.* Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.

KMK/Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2021). *Lehren und Lernen in der digitalen Welt. Die ergänzende Empfehlung zur Strategie „Bildung in der digitalen Welt“*. Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.

Lindner, A., & Berges, M. (2020): Can you explain AI to me? Teachers' pre-concepts about Artificial Intelligence. 2020 IEEE Frontiers in Education Conference (FIE), 1–9. https://doi.org/10.1109/FIE44824.2020.9274136

Ng, D.T.K., Leung, J.K.L., Chu, K.W.S., & Qiao, M.S. (2021), AI Literacy: Definition, Teaching, Evaluation and Ethical Issues. *Proceedings of the Association for Information Science and Technology*, *58*, 504–509. https://doi.org/10.1002/pra2.487

Ridsdale, C., Rothwell, J., Smit, M., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., Wuetherick, B., & Ali-Hassan, H. (2015). *Strategies and Best Practices for Data Literacy Education Knowledge.* Stifterverband.

Schmid, U. (2025). Grundkompetenzen im Bereich Künstliche Intelligenz (AI Literacy). In G. Brägger & H.-G. Rolff (Ed.), *Handbuch: Lernen mit digitalen Medien* (in print). Beltz.

Schmid, U., Gärtig-Daugs, A., Müller, L., & Werner, A. (2021). Grundkonzepte des Maschinellen Lernens für die Grundschule – Algorithmen, Biases, Generalisierungsfehler. *INFORMATIK 2021*, *Lecture Notes in Informatics*, 1611–1621. https://doi.org/10.18420/informatik2021-135

Schultz-Pernice, F., von Kotzebue, L., Franke, U., Ascherl, C., Hirner, C., Neuhaus, B., Ballis, A., Hauck-Thum, U., Aufleger, M., Romeike, R., Frederking, V., Krommer, A., Haider, M., Schworm, S., Kuhbandner, C., & Fischer, F. (2017): Forschungsgruppe Lehrerbildung Digitaler Campus Bayern: Kernkompetenzen von Lehrkräften für das Unterrichten in einer digitalisierten Welt. *merz – medien + erziehung, Zeitschrift für Medienpädagogik, 4*, 65–74. https://doi.org/10.21240/merz/2017.4.16

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What Should Every Child Know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence, 33*, 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

Vo, G. M., & Pancratz, N. (2023): Vorstellungen von Lehramtsstudierenden zu künstlicher Intelligenz. In L. Hellmig & M. Hennecke (Ed.), *INFOS 2023 – Informatikunterricht zwischen Aktualität und Zeitlosigkeit* (pp. 73–82). Gesellschaft für Informatik.

Wolff, A., Montaner, J. J. C., & Kortuem, G. (2016). Urban Data in the primary classroom: Bringing data literacy to the UK curriculum. *The Journal of Community Informatics, 12*(3), 57–82. https://doi.org/10.15353/joci.v12i3.3278