>TrainDL

>Teacher training for Data Literacy & Computer Science competences // Report Round 3 // **Deliverable**

4.7: Final Evaluation Report

train-dl.eu

Martin Reger, Evgenia Samoilova

University Potsdam

Content

List of Figures

List of Tables

Executive Summary

The TrainDL project provided policy recommendations for integrating the subjects of data literacy (DL) and artificial intelligence (AI) into professional programmes for CS teachers, STEAM, and primary teachers. This report synthesizes findings from three rounds of interventions and follow-ups, involving a total of 22 evaluated trainings across Germany, Austria, and Lithuania. The key outcomes include:

Training Impact and Competence Development: Pre- and post-training evaluations show at least some improvement in teachers' perceived competencies in DL and AI application in the classroom. The qualitative feedback consistently highlighted a strong appreciation for the practical components of the training but suggested a need for longer (longer than one day) and more in-depth trainings.

Feedback on Training Content and Format: The feedback was overwhelmingly positive. However, there was a recurrent theme across all teacher groups about the need for extending training duration to cover more comprehensive content and allow deeper engagement with complex tools and subjects.

Integration Challenges and Teacher Motivation: Follow-up surveys indicated a general willingness and a positive trend toward adopting DL and AI in teaching practices. Despite this, there were challenges related to the integration of DL and AI into curricula, particularly in adjusting content to diverse teaching contexts (for STEAM and primary) as well as overall overloaded curriculum. Teachers showed high motivation to continue learning about DL and AI, emphasizing their relevance to societal and educational advancements.

Recommendations for Policy and Practice: The evaluation suggests the need for policy adjustments to support the integration of DL and AI into educational frameworks.

This includes the development of comprehensive training modules that are iterative and extend over longer periods to accommodate deeper learning and application.

In conclusion, the TrainDL project has made significant strides in advancing teachers' capabilities and engagement in integrating DL and AI into their educational practices.

Introduction

The TrainDL project aims to provide policy recommendations for integrating the subjects of data literacy (DL) and artificial intelligence (AI) into teacher university education, as well as professional development programmes for teachers. To achieve this, the project adopts an iterative approach to design, deliver and evaluate teacher trainings. The project comprises three intervention rounds targeting computer science (CS) teachers as well as teachers from STEAM subjects at the primary and secondary levels. This report presents and discusses results of the three rounds of the intervention presented in the deliverables D4.4, D4.5, and D4.6. In addition to the data presented and discussed in these deliverables, this report discussed the data from the 2.5 day trainings that were conducted later (in November 2023 and January 2024) as well as follow-up data collected at least six months after the trainings, the later point of data collection did not allow us to include these data into the earlier deliverables. This report presents and discusses the results from the three rounds of the intervention, as detailed in Deliverables D4.4, D4.5, and D4.6. In addition to the data discussed in these documents, this report includes information from the 2.5-day trainings conducted in November 2023 and January 2024 and it incorporates the follow-up data collected at least six months after the trainings. Due to the timing of this later data collection, these data could not be included in the earlier deliverables.

Table 1 presents an overview of the evaluated training sessions, the evaluation instruments used as well as response rates. In total 22 trainings including 143 training

hours were evaluating, where the duration ranged from 90 minutes to 2.5 days. The evaluation was conducted on 376 participants, where 243 participants responded to both pre- and post-survey and 92 participants took part in the qualitative interviews.

The evaluation of the all three rounds of interventions focuses on teachers' capacity to integrate the acquired content on DL and AI into their teaching, as measured immediately after the trainings as well as with the follow-ups ca. 6 months after the trainings. While the reports for each intervention round were covering the round-specific findings in details, the main focus of this report is to present a consolidated summary across the three rounds. The summary is based on both the respective deliverables (D4.4, D4.5, and D4.6) as well as discussions of the results with the TrainDL team members.

To evaluate the training sessions, the following instruments were used:

- The evaluation survey administered before and immediately after the training as well as at least 6 months after the trainings;
- The DL and AI self-assessment and knowledge test, which includes both self-assessment and knowledge questions on DL and AI, administered before and immediately after the training;
- The semi-structured personal and online interviews administered right after the training as well as at least 6 months after the training.

Date	Location	Target group	Duration	Number of participants	Evaluation instruments and number of respondents/ percentage of the total number of participants
1st intervention round (in-service and pre-service CS)					
13.06.2022	Berlin, Germany	Pre- and in-service CS teachers at the	7 hours	24	• (Pre- and post-) survey: 20 (83%) • (Pre- and post-) DL and AI knowledge test: 16 (67%) • Interviews: 6 (25 %)

		secondary level			• Six-month follow-up survey: 7 (30%) • Six-month follow-up interviews: 2 (8%)
02-03.11.2022	Berlin, Germany	Pre-service CS teachers at the secondary level	3 hours	25	• (Post) survey: 18 (72%)
10.12.2022	Vilnius, Lithuania	In-service CS teachers at the secondary level	7 hours	21	• (Pre- and post-) survey: 16 (76%) • (Pre- and post-) DL and AI knowledge test: 14 (67%) • Interviews: 8 (38%) • Six-month follow-up survey: 21 (100%)
31.01.2023	Vienna, Austria	In-service CS teachers at the secondary level	7 hours	25	• (Pre- and post-) survey: 14 (56%) • (Pre- and post-) DL and AI knowledge test: 11 (44%) • Interviews: 6 (24%) • Six-month follow-up survey: 7 (28%) • Six-month follow-up interviews: 2 (8%)
25.-27.11.2022	Heidelberg, Germany	In-service CS teachers at the secondary level	2.5 days	19	• (Pre- and post-) survey: 18 (95 %) • Interviews: 5 (26 %)
2nd intervention round (in-service CS and STEAM teachers)					
09.03.2023	Berlin, Germany	In-service CS teachers at the secondary level	2 hours 15 minutes	12	• (Pre- and post-) survey: 17 (71%) • (Pre- and post-) DL and AI knowledge test: 8 (67%) • Interviews: 4 (33%) • Six-month follow-up survey: 5 (21%) • Six-month follow-up interviews: 1 (8%)
09.03.2023	Berlin, Germany	In-service CS teachers at the secondary level	2 hours 15 minutes	12	
17.05.2023	Graz, Austria	In-service teachers teaching STEAM subjects at	4 hours	8	• (Pre- and post-) survey: 8 (100%) • (Pre- and post-) DL and AI knowledge test: 8 (100%) • Interviews: 1 (13%) • Six-month follow-up survey: 2 (25%)

		the secondary level			• Six-month follow-up interviews: 1 (13%)
26.05.2023	Vilnius, Lithuania	In-service CS teachers teaching STEAM subjects at the secondary level	4 hours	12	• (Pre- and post-) survey: 12 (100%) • (Pre- and post-) DL and AI knowledge test: 10 (83%) • Interviews: 2 (17%) • Six-month follow-up survey: 8 (66%)
3rd intervention round (in-service CS and STEAM, and primary teachers)					
11.04.2023 18.04.2023 21.04.2023	Vilnius, Vilkaviškis region, Druskininkai, Lithuania	Primary teachers	4 hours	51[1]	• (Pre- and post-) survey: 30 (59%) • (Pre- and post-) DL and AI knowledge test: 30 (59%) • Interviews: 5 (10%) • Six-month follow-up survey: 51 (100%) • Six-month follow-up interviews: 2 (4%)
25.04.2023	Vienna, Austria	Primary teachers	3 hours	8	• (Pre- and post-) survey: 7 (88%) • (Pre- and post-) DL and AI knowledge test: 7 (88%) • Interviews: 2 (25%) • Six-month follow-up survey: 1 (13%) • Six-month follow-up interviews: 1 (13%)
13.09.2023	Düsseldorf, Germany	Primary teachers	4 hours	12	• (Pre- and post-) survey: 9 (75%) • (Pre- and post-) DL and AI knowledge test: 9 (75%) • Interviews: 4 (33%) • Six-month follow-up survey: 8 (67%)
10.10.2023	Vienna, Austria	Primary teachers	3 hours	11	• (Pre- and post-) survey: 8 (73%) • (Pre- and post-) DL and AI knowledge test: 8 (73%)

[1] While Freie Universität Berlin and OCG offered teachers trainings for which teachers had to register, the Vilnius university for the primary trainings reached out to the three schools directly and conducted their trainings in these schools. In addition to the 51 primary teachers (the main target group), 32 STEAM and 2 CS teachers participated in the training.

					• Interviews: 0 (no interviewees could be recruited) • Six-month follow-up survey: 8 (73%)
23.08.2023	Düsseldorf, Germany	STEAM teachers	4 hours	7[2]	• (Pre- and post-) survey: 4 (57%) • (Pre- and post-) DL and AI knowledge test: 4 (57%) • Interviews: 1 (14%)
05.09.2023	Berlin, Germany	STEAM teachers	1.5 hours	28[3]	• (Pre- and post-) survey: 26 (93%) • (Pre- and post-) DL and AI knowledge test: 26 (93%) • Interviews: 4 (14%)
23.10.2023	Graz, Austria	STEAM teachers	3 hours	12	• (Pre- and post-) survey: 8 (67%) • (Pre- and post-) DL and AI knowledge test: 8 (67%) • Interviews: 1 (8%) • Six-month follow-up survey: 8 (67%)
09.01.2024	Vilnius, Lithuania	STEAM teachers	4 hours	15	• (Pre- and post-) survey: 8 (53%) • (Pre- and post-) DL and AI knowledge test: 8 (53%) • Interviews: 1 (7%) • Six-month follow-up survey: 8 (53%)
18.09.2023	Berlin, Germany	CS teachers	4 hours	27	• (Pre- and post-) survey: 20 (74%) • (Pre- and post-) DL and AI knowledge test: 20 (74%) • Interviews: 4 (15%) • Six-month follow-up survey: 4 (15%)
24.-26.11.2023	Zeitz, Germany	CS teachers	2.5 days	24	• (Pre- and post-) survey: 18 (75%) • (Pre- and post-) DL and AI knowledge test: 18 (75%) • Interviews: 19 (79%) • Six-month follow-up survey: 8 (33%)
26.01.2024	Meißen, Germany	STEAM teachers	2.5 days	23	• (Pre- and post-) survey: 18 (78%) • (Pre- and post-) DL and AI knowledge test: 18 (78%) • Interviews: 20 (87%) • Six-month follow-up survey: 8 (35%)

[2] In addition, 2 CS and 3 primary teachers participated in the training.

[3] In addition, 1 CS and 2 primary teachers participated in the training

Table 1 Overview of the evaluated trainings of the three intervention rounds: date, location, target group, duration, evaluation instruments used, number of participants and response rates.

This report is structured into six sections. The first section outlines our research questions and evaluation methodology. The quantitative and qualitative findings are described in the sections 2 and 3 respectively. Sections 4, 5, and 6 wrap up the report, presenting a comprehensive summary of the salient findings across all trainings and discussing limitations.

1. Methodology

Deliverable 4.3 includes a detailed description of the methodology and research questions and hypotheses used for all three intervention cycles. The project uses an action research methodology (Baskerville and Wood-Harper 1996; Burns 2010), characterized by its iterative nature, involving multiple rounds of designing, implementation, observation, feedback, and reflection. In addition, Deliverables 4.4, 4.5, and 4.6 describe in details the data collection and data analysis methods that were used for the evaluations.

To evaluate the trainings, we followed the procedure outlined in Figure 1. To gain a more nuanced understanding of the trainings' impact, we employed a mixed methods approach following a concurrent nested design suggested by Creswell and Plano Clark (2018). This design allowed us to enrich and clarify our quantitative findings using qualitative data. The quantitative data were primarily used for the examination of the participants' characteristics and changes or lack thereof in teachers' perceived competences on how to use DL and AI in class as well as their understanding of these concepts introduced during the training. Additionally, qualitative interviews with the teachers provided additional insights into their experiences and perspectives on the training effectiveness. They also highlighted the teachers' expectations for future training content and identified barriers to DL and AI integration into the classroom.

Additionally, participants offered suggestions for policy changes that could better facilitate the inclusion of DL and AI into their teaching. The follow-up data collection captured the reported integration of DL and AI into the teaching. Detailed information on the instruments that were used is reported in deliverables 4.4, 4.5, and 4.6.

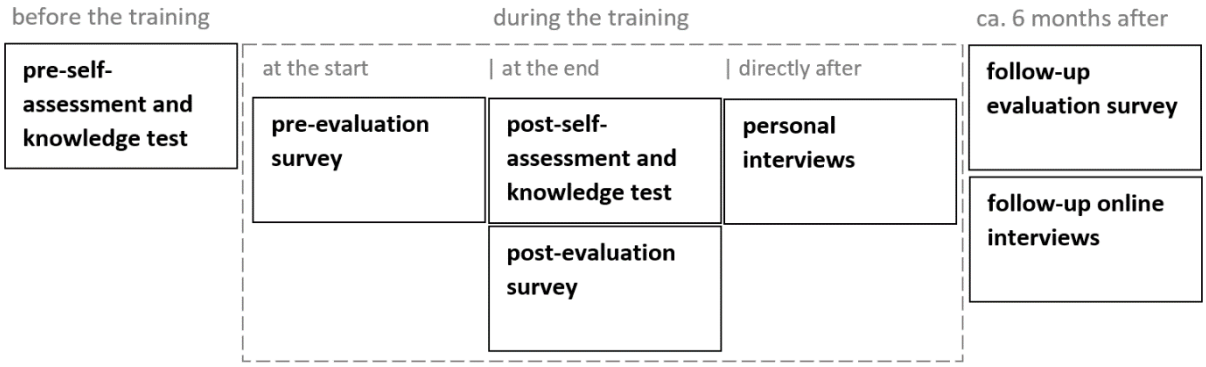

Figure 1 Overview of the evaluation process for each training in the second intervention round

To ensure the privacy of participants while still enabling the linkage of pre-, post- and follow-up datasets, participants were requested to create a unique pseudonymisation code, which they were required to enter or recreate during each subsequent round of data collection. This approach allowed for the protection of participants' privacy while maintaining the ability to connect and analyse the various datasets. Both the survey data and interview data were collected following informed consent, which included comprehensive information about anonymization, data storage, retention period, potential publication of anonymized data, and the option for participants to withdraw their consent and have their data deleted. The project did not require any personal data from the teachers, so no questions pertaining to personal information were included. Any personal information present in the interview transcript (such as place of work or names) was removed. Contacting potential respondents for the follow-up data collection is being done via local partners, eliminating the need to collect and store contact information.

The following definitions of DL and AI were used in the project and shared with the training participants, particularly in the evaluation surveys:

- DL is the ability to systematically handle data and consciously utilize and question them in the respective context. This includes the competences to collect, explore, manage, analyse, visualize, interpret, contextualize, evaluate, and apply data (Ridsdale et al. 2015).
- AI encompasses various technologies and methods that deal with the automation of intelligent behaviour such as decision-making, problem-solving and machine learning.

1.1 Quantitative data: research questions, instruments, and analysis

The quantitative data was primarily used to address the following questions:

1). *What is the effect of the designed DL and AI training on the ability of in-service CS and STEAM teachers to integrate DL and AI into their classes?*

2). *What is the effect of the designed DL and AI training on teachers' awareness of the significance that DL and AI have for their subjects?*

3). How does the level of motivation to acquire DL and AI competencies vary between computer science teachers and STEAM/primary teachers? (for the third round)

Integrating learned content after the training into teaching is a process that takes time. Therefore, the most suitable measures for such integration are the ones over an extended period. However, given the ongoing nature of the follow-up data collection and the anticipated low response rate inherent in contacting training participants six months post-training, the quantitative part of this report primarily focuses on perceived ability to integrate DL and AI into the classroom measured immediately after the training. Specifically, we looked at the following aspects with a focus on both pedagogical content knowledge and content knowledge:

- (pedagogical content knowledge) **teachers' perceived competences** on how to use DL and AI content in class, and
- (content knowledge) **teachers' understanding of DL and AI concepts** introduced in the trainings.

Additionally, for the teachers we have looked at the following aspects that can clarify and complement the main findings:

- teachers' feedback on the learned content and format of the training.

The quantitative data were also used to collect information on the socio-demographic characteristics of the participants. As random assignment to trainings was not possible, understanding the participants' characteristics was crucial for interpreting the findings and addressing potential selection bias. Each country's partner was tasked with internally advertising the training sessions, aided by local partners.

1.2 Qualitative data: research questions, instruments, and analysis

The research questions for the qualitative part of the evaluation included:

1). *How did participants perceive the training, and what suggestions do they have for enhancing the efficiency and effectiveness of future sessions?*
2). *How has the training influenced the integration of DL and AI into teaching, if at all?*
3). *How do participants evaluate the difficulties of conveying DL and AI concepts to students?*
4). *How can DL and AI be effectively integrated into the classroom, and what potential challenges could hinder this integration?*

The trainings were followed up by the two rounds of qualitative interviews right after the training using a semi-structured interview guide developed by the University

of Potsdam was used. The interviews were conducted in person by the evaluators or/and by the instructors with the help of the evaluators. The two 2.5 day trainings in the third round were evaluated with the help of focus groups, to ensure there is a sufficient amount of interviewees (there were only three longer trainings of 2.5 days and it was essential to collect sufficient qualitative data on them).

2. Quantitative results

2.2 Description of the sample

The average response rate for the three rounds of the pre-evaluation survey ranged from 80 % in the first and second round to 82 % in the third round, demonstrating a relatively high participant engagement. According to the pre-evaluation surveys, the trainings targeted teachers across various demographics and professional backgrounds in Germany, Austria, and Lithuania.

The age groups primarily ranged between 30-59 years across all training sessions and locations, reflecting country-specific statistics on age distribution of teachers. The demographics in the first CS round of interventions reveal varied gender representation across the locations with a generally lower proportion of female participants, especially in Vienna and Berlin's in-service sessions. The second round of interventions had similarly low ratio of CS female teachers, with more balanced gender distribution for the STEAM teachers. The third round of interventions once again reflected the gender disparities between CS and other subjects, especially the primary education. with female dominance in primary trainings and male dominance in CS trainings. Table 2 reports the most popular subjects that were reported by teachers for each target group (teachers usually had 2 subjects). For each target groups (computer science, STEAM, and primary teachers), mathematics was reported as the most frequent (second or first) subject.

Computer Science Teachers	STEAM teachers	Primary teachers
• CS 100% • Mathematics 47% • Physics 20%	• Mathematics 36% • Biology 31% • Languages 25% • CS 18% • Sports 16% • Chemistry 15%	• Mathematics 67% • Sports 59% • Languages 35% • Arts 27%

Table 2 Overview of the most popular subjects that were reported by teachers for each target group (teachers usually had 2 subjects).

2.3 Feedback on the learned content and format

The overall feedback for the learned content for all three rounds was very positive. 90% of the trained participants would recommend the workshops further[4]. The post-survey included a series of questions to assess participants' reactions to the suitability of the exercises presented in the training using an instructional strategy that allowed teachers to try out the activities designed for their students. Participants were asked to assess a level of perceived suitability of the activities for their teaching. As ca. 27 activities presented in the trainings varied, it is difficult to compare the findings across the trainings. Table 3 presents the top-rated activities for all of the three target groups. The activity on language models (which was not introduced in the primary trainings) were equally praised in both computer science and STEAM trainings. Activities that received a lot of mixed feedback included activities with Orange3, including work with regression.

Computer Science Teachers	STEAM teachers	Primary teachers
• Classical AI/Reinforcement learning: Beat the robot • Real world AI applications: AI-Bingo • Language Model: Grimms' fairy tales	• Generative AI: Schul-KI • Case study for image classification for Malaria, teachable machine • Language Model: Grimms' fairy tales	• Data literacy/intelligent search: Ligretto-like card game • Data mining with : Orange3 • Supervised learning/Decision trees: Good Bad Monkey

Table 3 Activities rated as most suitable activities for teaching for each target group

[4] albeit the survey item was introduced from the second round.

In addition, in the post-evaluation survey teachers were asked, "How much do you agree with the following statements?" and were given a scale from 1 ("not at all") to 6 ("definitely"):

- "I wish the training had lasted longer"
- “After the training, I am confident that I can independently implement the learned content in my teaching”

As Figure 2 demonstrates, on average all of the target groups tend to agree that they wish the training had lasted longer. The only exception is 2.5 days and to some extent 7 hours for STEAM. Interestingly, for CS the duration (for various lengths) was never enough and expressed a stronger agreement with longer trainings. For the primary trainings, due to real-life constraints only 1,5-4 hour trainings were offered, which did not allow us to compared different lengths for this target group.

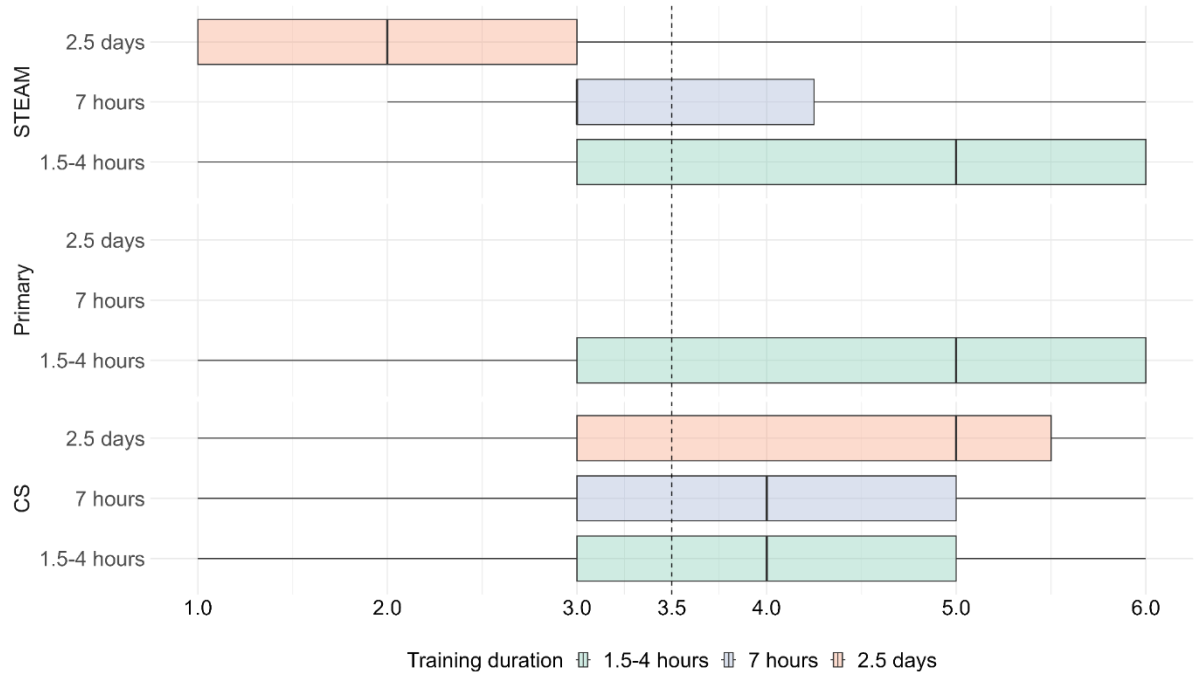

Figure 2 Post-survey evaluation survey, item - "I wish the training had lasted longer" for each target group

Figure 3 shows that compared to the primary and CS teachers; STEAM teachers report less agreement with the statement that they gathered enough competences to teach the learned content in class. Both primary and CS teachers report higher level of agreement compared to STEAM teachers, but for all of the target groups the level of agreement was rather moderate. The qualitative results elaborate on the challenges for the STEAM teachers to identify relevant for them DL and AI application areas.

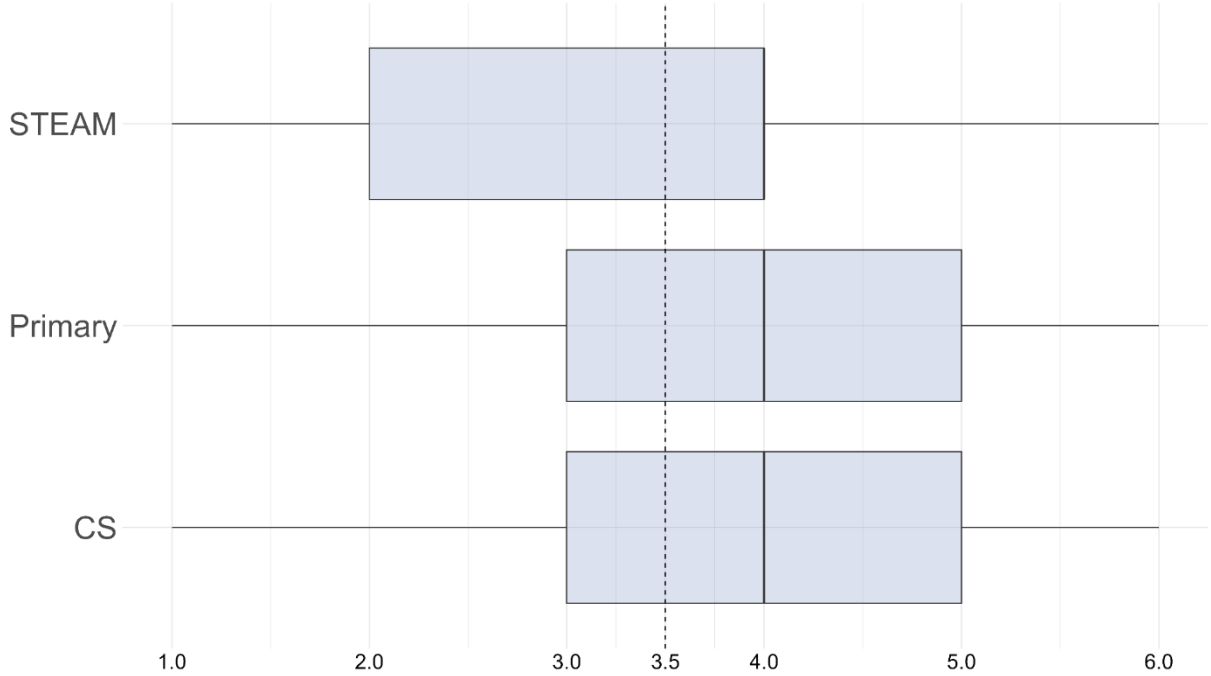

Figure 3 Post-survey evaluation survey, item - "Enough competences to teach the learned content in class" for each target group.

2.4 Perceived competences on how to use DL/AI in class

As demonstrated in Figure 4 and Figure 5, respondents reported an increase in the average self-reported level of competences in using DL and AI content for all of the target groups. The difference between the pre-and post-results was the most notable for the CS teachers, especially for the AI item, where interquartile ranges (IQRs) – the boxes representing the middle 50% of responses – appear narrower for the post-training data. This narrowing suggests a decrease in variability among participants'

responses after the training, indicating a more consistent level of perceived competence across the respondents. It's noteworthy that the pre-training self-assessed competence level of DL for CS teachers is lower than that for STEAM and Primary teachers, which may not be intuitive given the technical nature of CS education. The latter could suggest potential issues with the scaling or interpretation of the survey item (e.g., understanding of the item by participants with different depth of knowledge). The latter stresses the importance of comparing the pre- and post-differences within the target group and being cautious to compare the quantitative differences between the target groups.

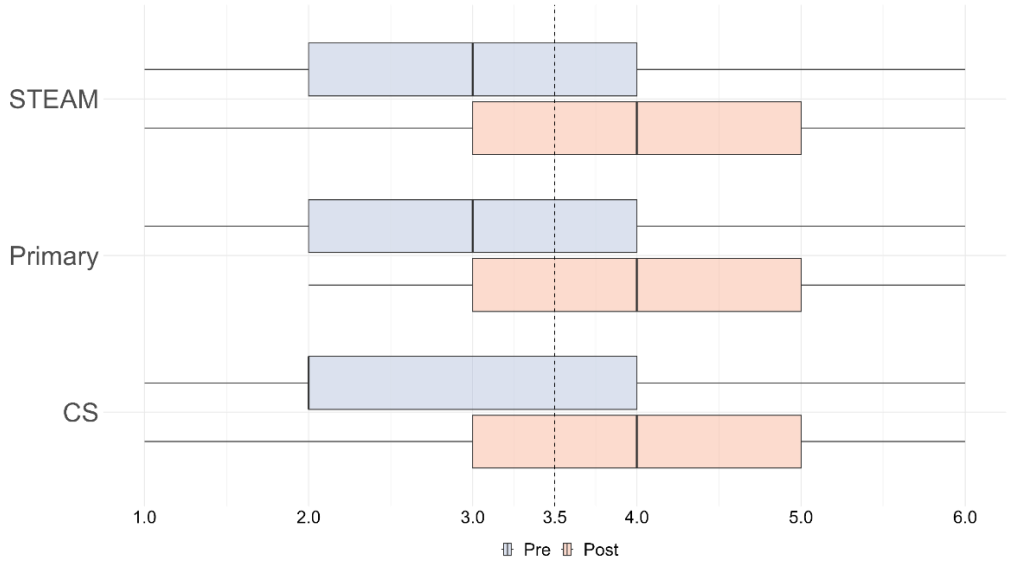

Figure 4 Boxplot comparison of pre- and post-results for the survey items on teachers' perceived competences to use DL content in class

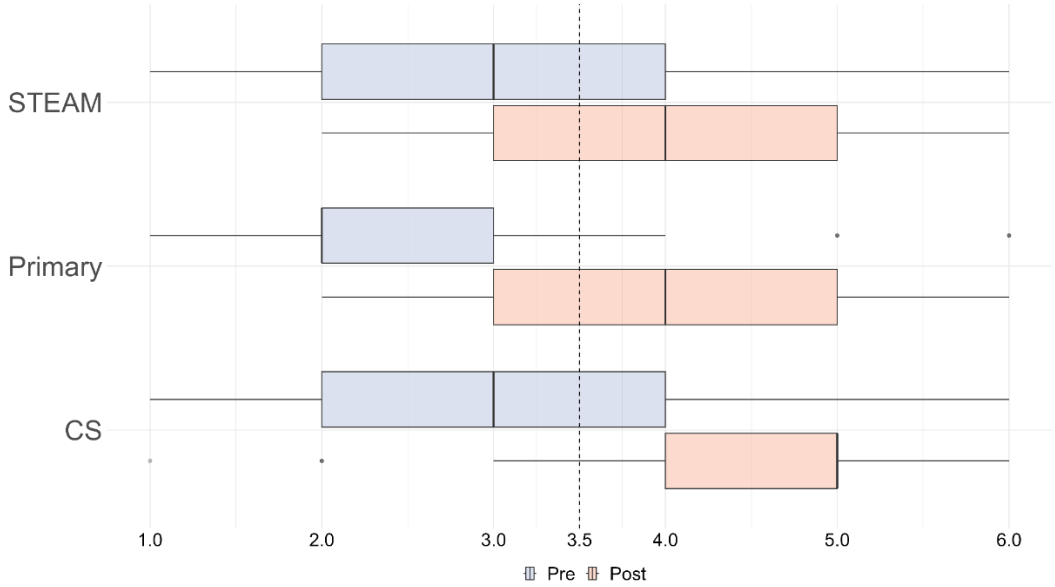

Figure 5 Boxplot comparison of pre- and post-results for the survey items on teachers' perceived competences to use AI content in class

2.5 Understanding of DL/AI concepts introduced in the training

All groups saw improvements in understanding basic concepts post-training, with the most pronounced pre- and post-differences for CS teachers, especially for the AI items. Questions on regression and using Orange3 were among the most difficult for all of the target groups. The purpose of the knowledge tests was to capture participants' basic comprehension of core concepts and definitions covered during the training sessions. The format and content of these assessments varied from one session to another, sometimes incorporating modifications to some of the questions in different iterations. Additionally, there was a report from a session in Lithuania indicating that some teachers expressed concerns regarding the accuracy of the results and collaborated on answering the questions. These challenges complicate the use of these results for evaluative purposes. However, the knowledge test still could prove beneficial for future training sessions serving a role of formative self-assessment.

2.6 Attitudes towards DL/AI and motivation to learn further

Immediately after the training, participants were asked about the societal importance of the subjects DL and AI for integrating them into the curriculum (see Fig. 6). Societal relevance was rated high by all the target groups, but the highest score was for AI in the group of CS, with a small spread, indicating a higher level of agreement compared to the DL item. Similar results (with AI being rated higher than DL) were reported for the item on "In the future, teaching DL/AI will provide added value to students," albeit for all of the target groups, not only for CS. Hence, as reported after the training, not all teachers saw DL and AI going hand in hand (not even CS teachers), where DL is being an integral part of AI.

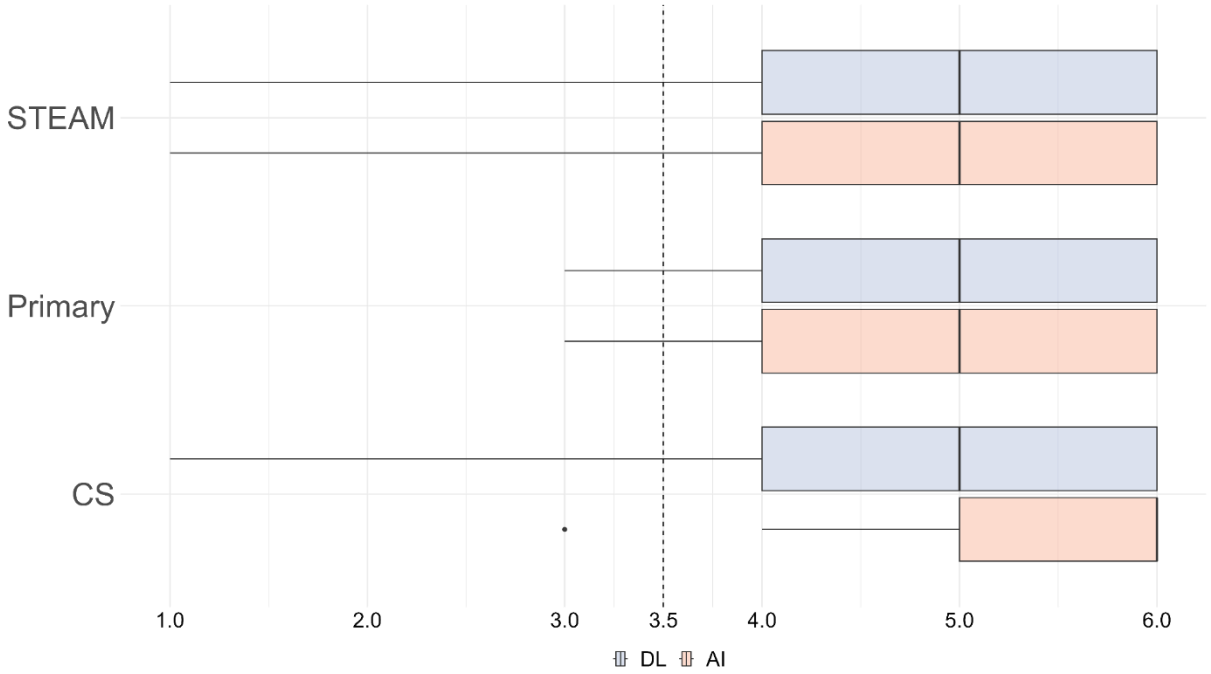

Figure 6 Boxplot comparison of pre- and post-results for the survey items "the topic of DL/AI is of enough societal importance to integrate it into the curriculum"

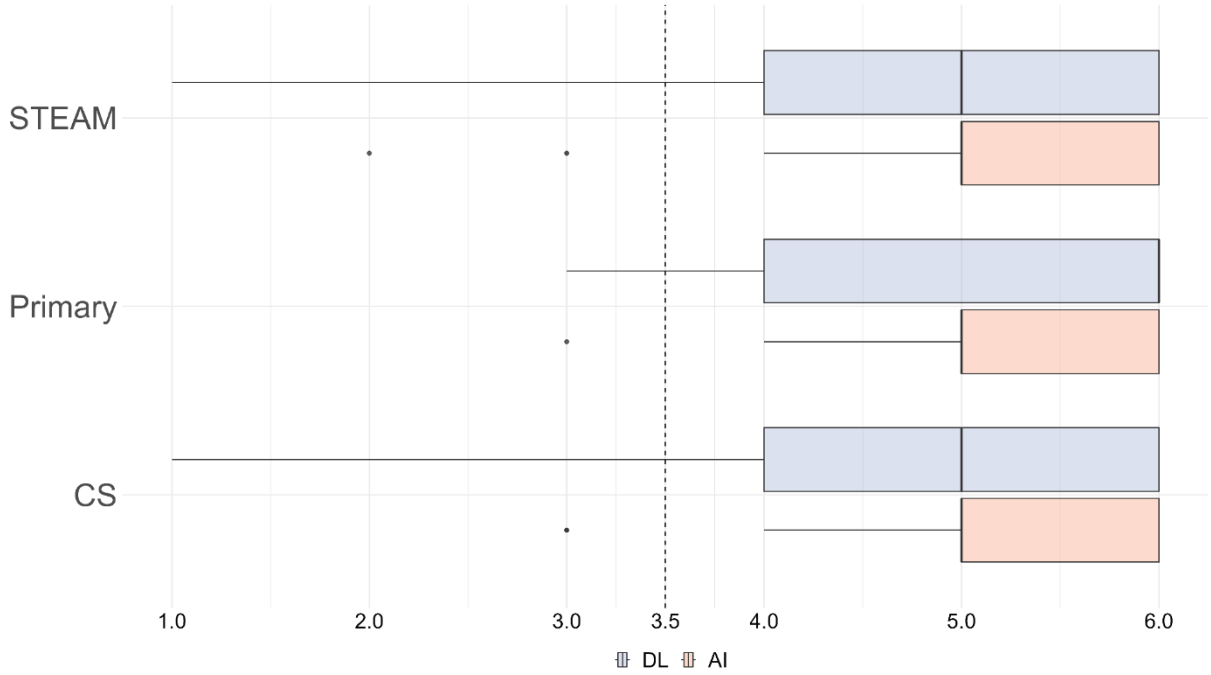

Figure 7 Boxplot comparison of pre- and post-results for the survey items "In the future, teaching DL/AI will provide added value to students"

For the question about motivation to learn further about DL and AI asked in the third round, all of the respondents except for one person indicated that they want to learn further about these topics.

2.7 Follow-up survey results

For the follow-up survey, conducted ca. 6 months after the training the response rate (with an exception of Lithuania in the first and third rounds) was very low. Therefore, the main focus for the follow-up data is on the qualitative interview, which allows to consider the aspect of the post-training integration of the learned content more in depth.

For the first round of CS interventions, the follow-up survey showed that in Germany ca. half of the participants reported teaching DL and AI after the training. In Austria, one third of the participants reported teaching DL and all apart form one participant reported teaching AI. In Lithuania, 14 out of 19 reported teaching DL and AI after the training. Similarly, in the follow-up survey for CS in the second round (CS

training in the second round was conducted only in Germany) showed that all but one respondent integrated DL and AI in their teaching.

For the STEAM in the second round, the sample sizes were very small. In Austria, one out of the two respondents reported the actual integration of at least one topic of the learnt content in class. For Lithuania, the numbers were seven out of eight respondents.

For the primary interventions, the Austrian sample included unfortunately only one participant, who reported teaching at least one topics in class. For Lithuania, the reported integration of at least one topic included 45 out of 50 respondents.

2.8 Summary of the quantitative results

The average response rate for the three rounds of the pre-evaluation survey ranged from 80 % to 82 % in the third round, demonstrating a relatively high participant engagement. The biggest challenge was to obtain a sufficient sample size for the merged pre- and post-survey: some participants did not fill out the post-evaluation survey. The response rate for both pre- and post-evaluation survey for the first round, where results were merged constituted between 44-67 %. For the second and third round we tried to tackle this problem by lowering the difficulty threshold for participants: reducing the amount of survey items focusing on the most important ones, combining the evaluation survey with the knowledge test into one instrument, so that students did not have to switch between the instruments and enter their pseudonymisation codes, explaining more why evaluation is important before announcing it. The response rate for the merge pre- and post-evaluation sample in the second and third round was higher than 70% in 10 out of 15 trainings. For 5 trainings, the response rate constituted 59-76 %, albeit most of them had a very small number of participants.

Compared to the CS trainings, STEAM and primary had a lower number of participants. The evaluated primary trainings in Lithuania attracted a relatively large number of participants, which might be due to them actively reaching out and traveling directly to the schools, where the trainings were conducted.

The overall feedback on the training content and format was positive. on average all of the target groups tend to agree that they wish the training had lasted longer. The only exception is 2.5 days and to some extent 7 hours for STEAM. The activity on language models (which was not introduced in the primary trainings) were equally praised in both computer science and STEAM trainings. Activities that received a lot of mixed feedback included activities on DL with Orange3, including work with regression. Similarly, the knowledge questions on Orange3 and DL seem to be more difficult as well as the reported attitudes indicate teachers (even CS) might not have a clear understanding about the role of DL in teaching AI.

Although the overall feedback was positive, teachers report a rathe moderate level of agreement with the statement that they gathered enough competences to teach the learned content in class. The level of agreement is especially low for STEAM teachers.

High societal relevance and value of teaching DL and AI in class as well as a uniform willingness to learn more about DL and AI for all of the target groups indicate a strong potential for integrating these topics into both teacher training and framework curriculum.

The follow-up data collected ca. 6 months after the training suggests a general willingness and positive trend towards adopting DL and AI in teaching, but the small sample sizes in some groups (primary in Austria) notably limit the generalizability of these findings. Therefore, the follow-up qualitative interviews appear here to be central for more understanding of the post-training integration.

3. Qualitative results

The analysis of the qualitative data is structured as following: first, the summary of the interviews with the teachers from the CS, STEAM, and primary trainings (already reported in details in D4.4, 4.5, and 4.6) is presented. It is than followed up by a more detailed description of the qualitative data for the three 2.5 day trainings: CS training 25. until 27.11.22 in Germany (Heidelberg) with five persons interviewed; CS training 24.11.23 until 26.11.23 in Germany (Zeitz), with 19 persons interviewed; STEAM training 26.01.24 until 28.01.24 in Germany (Meißen), with 20 persons interviewed.

The themes below describe key patters that repeatedly surfaced in the data, which was gathered based on the questions posed during the interview conducted immediately after the training. The labels on the right correspond to categories that were assigned inductively throughout the qualitative analysis process.

3.2 Training feedback

CS in all rounds, length 90 minutes – 7 hours:

The qualitative interviews from three rounds of CS teacher training reveal a consistent appreciation for the introduction to AI and DL, though participants noted challenges with the depth and complexity of certain topics. Mixed feedback was received on tools like Orange3, with some teachers finding it overly complex, while others appreciated its potential for classroom application. Across all rounds, the training design was generally praised for aligning well with participants' existing knowledge levels, and the blend of theoretical and practical components was well-received. However, feedback on the duration and structure of the sessions varied, with some teachers feeling the training was too short to cover all intended materials and exercises, while others considered it too lengthy. Suggestions for future trainings included more in-depth content, longer series of trainings, and the provision of ready-to-use, grade-appropriate materials that could better facilitate the integration of AI and DL into the curriculum.

STEAM 2-3 rounds, length 1.5-4 hours:

Overall, while the trainings were positively received in terms of their structure and initial impact, there was a call for adjustments in duration and content depth to better serve the diverse educational backgrounds and subject expertise of STEAM teachers. In the second round, the training was perceived as moderately difficult emphasizing the value of extending the training into a series of sessions spread over weeks or months. Such a format was seen as beneficial for delving deeper into specific programs and enhancing practical application skills, thereby better preparing teachers to convey complex topics to students. In the third round, the content was assessed as not fully applicable to STEAM teaching context. Both rounds shared common suggestions for improvement, focusing on the need for longer training sessions and the inclusion of more hands-on exercises to enhance practical understanding.

Primary 3rd round, length 3-4 hours:

The primary training design was well-aligned with participants' prior knowledge levels. Overall, the trainings were positively received. The length was considered appropriate in relation to the content. However, more trainings were welcomed, and there was a desire for longer sessions, although some participants found whole-day training too lengthy. It was suggested that teachers should be able to specify their knowledge in advance to ensure suitability of the training. The theoretical inputs and practical tasks, such as Ligretto, were praised for making topics clearer. However, participants expressed a wish for a more hands-on approach instead of primarily engaging with worksheets. Structurally, there was feedback about too much thematic jumping back and forth, so the suggestion was having one block each for DL and AI.

In terms of materials, there was a positive view, noting that they are practical and implementable in the classroom. However, participants suggested including more background information and establishing a stronger connection to the every-day life of students.

CS Germany 1st round, length 2.5 days:

Design of training: Well aligned with prior knowledge levels (DL: no experience)

Some participants indicated that the training overall corresponded well their prior knowledge, like: *"The basics were well explained but not too detailed, and that was completely sufficient for me. I can imagine that it was a bit lacking for people who have no idea about it at all. But for me, it was just right"* (Participant D - 25.11.22 until 27.11.22, Heidelberg).

Overall feedback: Positive; challenging in a good way

The training was classified as very good and beneficial: *"I found the event overall really successful because it addressed, repeated, and presented many things on different levels"* (Participant C - 25.11.22 until 27.11.22, Heidelberg). Furthermore, it was challenging for the participants in a positive sense*: "[...] since I am essentially only a lower secondary school teacher for computer science and have not studied it formally, but rather learned it in a self-taught manner, I would say that the difficulty level was already high for me, quite significant for me. But you can handle it. I mean, we all have problem-solving skills ourselves, and do we not expect that from our students? Therefore, I think it's okay to be challenged or pushed a bit beyond the usual level in training"* (Participant A - 25.11.22 until 27.11.22, Heidelberg).

Overall feedback: Positive – Bridge between AI and DL became clear through training

What also became clear through the training is that data (collection) is very important for AI, for example: *"[...] data determine how an AI behaves, and then of course it's also about society, which data do I use. What does it mean when I have data that already perhaps have certain aspects, biases, boundaries. Then that is automatically reflected in the AI. So, you can really draw connections [between DL an AI]"* (Participant C - 25.11.22 until 27.11.22, Heidelberg).

In general, it was said that the format of the weekend workshop was very successful.

Format: Execution of the workshop on the weekend (very) good

Positive feedback regarding the content highlighted that inspirations for teaching could be conveyed (for example, through the museum exhibition "I am AI"). Solid foundations for classroom use were established, especially concerning the initial understanding of AI. Therefore, the training served as a good didactic introduction for teachers to DL and AI, also reinforcing technical terms for participants already familiar with such concepts.

Format – Content: Positive: Good inspiration for future teaching, exercises/content good foundation (e.g., for teaching students), Good introduction, theoretical conceptualizations / concepts / terms strengthened

Furthermore, it was positively praised that concrete examples, respectively exercises were completed in the training. It was great that they were practically applied, allowing teachers to try them out. As a result, new or never-before-tried tools such as Jupyter Notebooks or Orange3 could be tried out, which was an exciting experience for the participants.

Format – Content: Positive: Concrete examples / exercises, practical tasks well liked, learned new exercises / tools

Regarding specifically mentioned exercises, the guest lectures were praised as informative and engaging, among other things, *"[...] because one must also narrate AI in schools a bit lively, and when one has information about what's happening in practice, on the front lines, we as teachers never get to experience that. It's enriching to bring the topic of AI a bit closer to the students. So not only from the theoretical side"* (Participant C – 25.11.22 until 27.11.22, Heidelberg). The museum tour for data collection and subsequent application in the AI context was praised because the teachers could try to collect data themselves and it was similar to a project that could be carried out with students, thus providing teachers with suggestions for teaching. This could allow

Format – Content: Positive: Guest input from business (presentation/lecture) (very) good, guided tour in the museum / "I Am AI" exhibition

for a whole process from data collection to analysis in the classroom, resulting in comprehensible results.

Orange3 was praised because it makes individual process steps clear in working with data, and therefore could be pedagogically useful. Additionally, the "Good Monkey – bad monkey" game received very positive feedback.

Format – Content: Positive: Orange3: good / very good / interesting, Good monkey - bad monkey

The length of the training was praised. With the 2.5-day-duration, it is guaranteed that the individuals who are truly interested in these topics are reached, and these individuals can then also act as multipliers. Additionally, the two and a half days provide (for CS teachers) an appropriate framework to familiarise oneself with new and complex topics. Moreover, there are plenty of opportunities to exchange ideas with colleagues.

Format – Length: Positive: Workshop over numerous days = Good / very good, reach only those who are interested, participants can act as multipliers, familiarisation with topics possible, more exchange with colleagues

The transition between theoretical content (lectures) and active exercises, was perceived as very beneficial.

Format – Structure: Positive: Good mixture between theory and exercises

The level of interaction was evaluated as *"[...] great, it fit well and there was always the opportunity to ask questions and learn something additional"* (Participant D - 25.11.22 until 27.11.22, Heidelberg). The individuals who conducted the event communicated the topics well.

Other factors: Interaction / communication: Positive - (Very) good (General statement), participants always could ask questions, (very) good speakers

The staff-to-participant ratio was praised, as well as the organization of the event. It was very fruitful that teachers from different federal states were present, allowing for collaborative work and exchange, not only among teachers but also with people from the university sector.

Other factors: Positive - Staff-to-participant ratio adequate, (very) good organisation, new perspectives through teachers from other federal states, exchange with other teachers (very) good / interesting

When it comes to discussing expressed ideas, instructions for installing programs such as Orange3, which are needed for the training, should be sent out to the participants earlier, for example, one week before the start of the training. Also, since the workshop is very cost-intensive, it should be considered how to make the training more accessible to a broader audience.

Ideas for further training: Training-specific (Overall training) – Send instructions for installation (in advance) earlier, consider how to offer a workshop broadly / cost-effectively

One approach that can be incorporated into the training is *"[...] the concept of flipped classroom [...]. On one hand, it's very helpful if you can outsource certain content and say you can somehow [make] a preliminary video [...], so to speak, so that you know you already have a certain foundation or a basis upon which everyone can build upon"* (Participant A - 25.11.22 until 27.11.22, Heidelberg).

Ideas for further training: Training-specific (Overall training) – Flipped classroom approach

General ideas were also proposed, like the establishment of *"[...] 'micro-trainings,' meaning trainings conducted by colleagues for colleagues. So, in every school, there are colleagues who are already more deeply involved in the topic, who are more interested, and who can pass on their knowledge. And I believe strongly in this because then, in a familiar environment where one also knows what is technically possible, one can actually receive further training from colleagues whom one knows personally. So to speak, in a small-scale training session, they can certainly serve as multipliers [teachers] and may have been trained externally at first, but this could ensure that a large [number of] colleagues could perhaps be trained in a relatively short time, by having only individual colleagues perhaps attend external training and then imparting it to their colleagues"* (Participant A - 25.11.22 until 27.11.22, Heidelberg). Furthermore, efforts could be made to bring in experts from the university sector who can break down data processing and AI for school purposes.

Ideas for further training: General – Teachers train teachers at own school, experts from universities who can break down content for school

Ideas for content were also highlighted in the interviews. For example, as an icebreaker at the beginning, participants could be asked to share which AI applications they use in their daily (work) lives. This could emphasize the importance of AI in the school context right from the start. Thematically, there was a sense that the training was very focused on DL, and it was suggested that more emphasis could have been placed on AI. Additionally, the various types of AI should be introduced in more detail, in order to create a better understanding of the concepts among all participants for the remainder of the training.

Ideas for further training: Format – Content: Teachers tell about AI applications (as icebreaker), more emphasis on AI, better differentiation (of types) of AI

When it comes to specific exercises, one suggestion was to break down complex tools into understandable units. Although the guest lectures were interesting, it would be worth reflecting on whether the considerable amount of time they took up would have been better spent elsewhere. Additionally, it was suggested to invite experts from the university sector, *"[...] because AI or data literacy, or whatever you want to call it, is already taught at universities. [...]. And I would have expected more of that perspective as well, because, of course, when we teach university students who have decided to study, we can teach at a higher level than in secondary school levels one or two. But then there should already be a lot of experience in how to convey this content"* (Participant B - 25.11.22 until 27.11.22, Heidelberg).

Ideas for further training: Format – Content: Breaking down complex tools into understandable units, time from business experts (presentations/guest lectures) might be better used for other content + input from university instead of from business representatives

Furthermore, there was a desire to introduce an exercise on image recognition because students find this topic very exciting. For the project of the guided tour and the exhibition 'I am AI' with data collection, more time was desired, or a possible deepening in another event with the same participant group. Another participant said the

Ideas for further training: Format – Content: images (e.g., manipulation, recognition), 'I am AI' exhibition / guided tour, case study: more time or less time

opposite, namely, that this exercise should have taken less time in retrospect.

For one-day trainings, it was stated that they are "[...] usually short, so of course you get something out of it, [but] [...] it's more like a motivational event" (Participant C - 25.11.22 until 27.11.22, Heidelberg). Half-day events such as three-hour trainings would also be too short for deeper knowledge dissemination.

Ideas for further training: Format – Length: Disadvantages of full-day trainings (e.g., 7-hour trainings) or half-day-trainings: Deeper knowledge dissemination not possible

A greater emphasis on practical teaching methods was desired, for example, through materials that could be immediately implemented the next day.

Ideas for further training: Materials: Concrete implementation plans

CS Germany 3rd round, length 2.5 days:

The training provided the correct terms, concepts, and definitions necessary to address AI in the classroom, thus also conveying fundamental knowledge. In some cases, while some basics were already known, the training also imparted additional knowledge.

Overall feedback: Positive – Training has provided input (e.g.,for application in classroom), foundational knowledge through training aquired, good / very good, because knowledge was disseminated

Additionally, the training also contributed to structuring existing knowledge. For example, one participant already possessed theoretical knowledge but was surprised by how easy it is in practice to predict the price of a pizza using AI and linear regression. Furthermore, it helped to contextualize prior knowledge of data competencies in a more specific context.

Overall feedback: Positive – Training has helped to structure (existing) knowledge, surprised how easy linear regression is, data competence / DL was brought into new context

In the areas of DL and AI, according to the participants, knowledge enhancements have occurred, such as: *"My prior knowledge in the field*

Overall feedback: Positive - Knowledge in DL and AI has increased

of data competency has increased because one has seen what it takes to generate a usable dataset" (Group A – 24.11.23 until 26.11.23, Zeitz).

Overall feedback: Neutral - Same / similar knowledge (for AI, or DL and AI) as before

However, some participants also said that their knowledge of DL and AI was roughly the same before and after the workshop. In some cases, this question was specified regarding moderate knowledge both before and after. Another participant said that their knowledge had only increased slightly because the content on AI in the training was not sufficiently extensive.

Format – Content: Positive: Good / very good / interesting (general statement), training and practical exercises (concrete examples) good inspiration and basis for future teaching, societal-relevant questions about AI

The content of the training was perceived as very interesting. More workshops like this would be desirable. Several ideas and information could be conveyed for use in the classroom, such as: *"I still feel confident in integrating Data Literacy. I may have received individual insights into certain ideas and projects. For example, I am currently planning a project and have received certain impulses from other colleagues on how to implement it and what other ideas there are. This has helped me because exchanging ideas within the school, due to the small department, is often not so easy"* (Group B – 24.11.23 until 26.11.23, Zeitz). The practical and specific examples, respectively exercises were also a good preparation for future classroom use of the topics. Except for the pizza project (data workshop), one participant said that all other exercises were relevant and good. For one participant, the societal aspects of AI were particularly fascinating, or more fascinating than some of the practical exercises.

Format – Content: Positive: Practical tasks well liked

The hands-on approach of the training, respectively the exercises, was very appealing. One interviewee indicated: *"It gave me a good feeling that I can handle what I learned. It wasn't just passive teaching; we also applied our knowledge. I think that's good"* (Group D – 24.11.23 until

26.11.23, Zeitz). Some of the exercises would be interesting for students. There were tools that were known by name, but practical experience with them had been lacking until now. But also, new tools could be learned through the practical exercises.

Format – Content: Positive: Pizza project (data workshop) with preceding data collection, Orange3, Abalone

When specific exercises were mentioned, it shows that the pizza project and the corresponding data workshop on Saturday were positively assessed; it was demonstrated there how important the data collection conducted before this project was. Working with Orange3, such as in the context of the Abalone exercise, which was also classified as instructive and relevant, especially considering animal welfare considerations, was also *"[…] beneficial for evaluating a linear regression, planning a dataset, acquiring the data, and then cleaning it"* (Group A - 24.11.23 until 26.11.23, Zeitz). Furthermore, the Abalone exercise was positively highlighted because it illustrated a project process from start to finish.

Format – Content: Positive: Development of an app (part of 2nd data workshop), connection of exercises to food very good, food inspection, presentation / lecture very good / good

As part of the data workshop on Sunday, the guided development of a price predictor app was highlighted as meaningful, useful, and well-received by several participants. Generally, the food-related aspect of many exercises was perceived as appealing and interesting approach: This included the previously mentioned pizza project (data workshop on Saturday), but also refers, for example, to the guest lecture about food and AI and the food inspection exercise (algorithmic bias and image recognition); the latter was *liked "[…] because it is illustrative, one gets results directly, and one can use it directly in class"* (Group A - 24.11.23 until 26.11.23, Zeitz).

Format – Content: Negative: Pizza project (data workshop)

Alongside these mixed reactions, also negative criticism regarding specific exercises was mostly focused on the pizza project. A few

times, it was explicitly identified as the weakest task. Questions were unclear or incomplete on the accompanying worksheets. The objective and context were also unclear or irrelevant, for example: *"I found the question of how much a pizza costs when we open our own shop to be poor. It has little relevance to students' real lives. In contrast, it is important that we don't have to kill an abalone to see how old it is. The product and relevance are clear to me in that case. I didn't have that with the pizza project"* (Group A - 24.11.23 until 26.11.23, Zeitz). Additionally, as mentioned earlier, too much time was spent on entering data. The data collection itself was also questioned or not goal-oriented, as one could have researched the prices on the internet without having to visit pizzerias for that purpose. Furthermore, no results were achieved, meaning no pizza was created with the help of AI.

Format – Content: Negative: Generative Cocktail Producer, data life cycle

For the task of Generative Cocktail Producer, the creation of a weighted graph was not seen as an added value. Furthermore, no results were achieved. The training lacked reference to the data life cycle.

Format – Content: Negative: Alignment of training content to number of class hours not evident, initial uncertainty in dealing with (unknown) tools / concepts, not enough time for (some) exercises, time was wasted on certain content, differentiation unclear: didactic and factual knowledge

It was unclear how long it would take to teach the content in class provided in the training. Additionally, the definition of terms/concepts was not clear, for example: *"Problematic for me was the initial use of these tools because I did not have a complete understanding of the terminology. This not only led to misunderstandings but also to ignorance about how to interpret the term precisely"* (Group C - 24.11.23 until 26.11.23, Zeitz). Furthermore, time was perceived as too limited to take part in all exercises. Additionally, too much time was spent on some exercises, although there was no usable outcome. The separation between the factual knowledge that is needed and the

knowledge that is needed in practice to convey the topics in the classroom was unclear.

The unplugged teaching material on language models would be ready and immediately usable, for example: *"I liked the unplugged material for the language model because it could be tried out by hand. This puts you on the right side with the students. They can start right away and learn how it works"* (Group A - 24.11.23 until 26.11.23, Zeitz). The immediate use of the material was related to, for instance, the food control material or establishing an elective course at the school of one of the interviewees.

Material: Positive: Materials practical and (instantly) usable for classroom, unplugged material for LLM praised

In relation to other factors, the exchange facilitated by the training with other teachers was often praised; specifically, the contact with colleagues from other states within Germany was also mentioned which opened new perspectives.

Other factors: Positive: Exchange with other teachers (very) good / interesting, new perspectives through teachers from other federal

Following that, ideas for further training were expressed, including the desire for even more space for collegial exchange, for example, *"[...] regarding the specific conducting of a classroom lesson or examples of how to approach it in class. Why was the concept developed this way, and what do other colleagues say about it? So that I may develop an idea for the first or a subsequent lesson. How do I introduce the topic? How do I proceed with it?"* (Group C - 24.11.23 until 26.11.23, Zeitz).

Ideas for further training: Training-specific (Overall training) – More communication with colleagues

There was a desire for another workshop on the Saturday (25.11.23) following the pizza project (data workshop), as this project did not work out so well. Although the idea itself for this exercise was appreciated, other questions could have been addressed, such as in-

Ideas for further training: Format – Content: One more slot for a worksho, pizza project (data workshop), practical / real world application of topics / exercises

troducing a graph showing which ingredients are combined how often. For such or other tasks, more time would have been needed, time that could have been saved during data entry, which did not add any value for the participants. Additionally, instead of going to pizzerias, one could have specified fruits and vegetables in the supermarket. Exercises should be closer to the students' everyday life context; for example, one could compare hardware properties of mobile phones and their effects on prices (What price do you have to pay to get which features?). That would be more practical-oriented, as one can conduct internet research instead of visiting places on-site (as in the pizza project).

Ideas for further training: Format – Content: Differentiation for: didactic and factual knowledge

A differentiation was requested *"[...] between what I need to expand my knowledge enough to teach it [topics of the training] and what I need in practical teaching to pass on this knowledge. These are two completely different things for me. Either I didn't understand it or it was mixed up here. We did many exercises that the students should carry out. This also coached me, but I would have liked a separation of these aspects"* (Group C - 24.11.23 until 26.11.23, Zeitz).

Ideas for further training: Materials: More material connected to students' everyday life, Orange3: disable unused functions, so that it is easier to handle, provide information about plugged tools / alternatives

Also, different materials/exercises were desired, *"[...] something closer to the everyday lives of children than the Abalone"* (Group C - 24.11.23 until 26.11.23, Zeitz). For Orange3, reducing complexity would be sensible, so that one can deactivate those options/widgets that are not needed in training. Additionally, web tools/plug-in alternatives for teaching were asked for, which can be accessed without the need to log in/register—however, such tools are not currently known

STEAM Germany 3rd round, length 2.5 days:

The training was well received as it provided input that can be used in the classroom. Fundamental knowledge about DL and AI could be imparted to the participants, as well as how certain related processes work (e.g., data cycle).

Overall feedback: Positive – Training has provided input for application in classroom), foundational knowledge aquired, some things examined in more detail

The training provided *"good, general input"* (Group A – 26.01.24 until 28.01.24, Meißen). Scientific references were of interest. The training was successful as questions brought by the participants to the workshop were answered during the weekend, thus knowledge was imparted.

Overall feedback: Positive – (Very) successful / interesting / good / valuable (General), scientific aspects (very) interesting, good / very good, because knowledge was disseminated

When asked about the comparison of knowledge before and after the training, there were only statements attesting to a gain in knowledge after the training. For example, regarding DL: *"But the way it was designed here, with examples from Grimm's Fairy Tales [exercise] and also yesterday's exercise with the abalones, I found that very instructive. This is how one actually understood what this data collection and data analysis means"* (Group B – 26.01.24 until 28.01.24, Meißen). But also in AI, there was a gain in knowledge, such as: *"The training helped me in the sense that I now fundamentally understand what AI means or what the range of meanings is. I now realize that it's not something super new, but something evolved"* (Group B – 24.01.26 until 26.01.24, Meißen).

Overall feedback: Positive - Knowledge in DL and AI has increased

A neutral assessment was given because after the training, some questions remained unanswered.

Overall feedback: Neutral - Some questions have remained unanswered

When asked about knowledge acquisition, it was expressed that one still doesn't feel confident enough. Additionally, the transferability of the training content was questioned (despite fundamental competencies being conveyed.). Although the training was designed

Overall feedback: Negative - Not (completely) confident in own competence post-training, transferabilty of training's content not clearly seen, for non-CS teachers /

for STEAM teachers, it was perceived by one respondent as *"[...] a bit too computer science-heavy. For someone who has little knowledge of linear regression or mathematics, it's difficult to keep up. That's then unfavorable"* (Group B - 26.01.24 until 28.01.24, Meißen).

mathematicians sometimes hard to follow

In terms of content, it was positively noted that on the first day, there was initially an explanation of what AI is. Additionally, it was noted that after the workshop, it was clear what AI is. The training provided a good introduction/foundation to the subject area of DL and AI.

Format – Content: Positive: Introduction what AI is on first day (very) good / helpful, AI became clearer through input and exercises, good introduction

Positively highlighted was that content was secured with a flip chart / display board. Content of the training would also be transferable to the classroom. There was also info about applications in research, which was interesting. The exchange in subject groups, which took place at various points during the training, was perceived as fruitful.

Format – Content: Positive: securing the results (e.g., with a flip chart), content transferable into teaching, learned about applications in research, discipline-specific/subject input/discussions with colleagues

When the discussion in the interviews shifted more towards the activities that were present in the training, it was underlined, that practical engagement with the tools and exercises that can be used in the classroom was very helpful. It was emphasized that unplugged exercises generally have the advantage of being independent of technology or not being disrupted by technology when applied in the classroom.

Format – Content: Positive: Practical tasks well liked, exercises / tools had connection to teaching in the classroom, analogue games / unplugged exercises (very) good

To address specific activities, the "Teaching Workshop: Idea Sketches for AI and Data in Subject Teaching" was generally considered helpful and important, as well as "AI and Data: What Skills do Teachers Need?" to develop suggestions for teaching. "The activity

Format – Content: Positive: Teaching workshop, AI & Data, AI-Quiz, food inspection, Abalone, Grimm's New Fairy Tales, Miro Board (very) good

"AI Quiz" was helpful in gaining understanding about AI. The exercises Food Inspection, Abalone, and Grimm's New Fairy Tales received repeated general approval. The latter also contributed to the understanding of AI: *"I found especially intriguing the background information about these games we played, breaking down the sentence, how to build the AI model. But also how complex it is and how much data is inputted for such an AI to even function"* (Group D – 26.01.24 until 28.01.24, Meißen). Also, the Good Monkey-Bad Monkey-Game was praised, as well as the exercises focusing on Generative AI. The access to or working with a Miro board was positively acknowledged.

Format – Content: Positive: Guest input (presentation/lecture) (very) good

Speaking of theoretical inputs, the expert lectures were also appreciated, focusing *"[...] mainly on how we can enrich our knowledge through dealing with data and AI"* (Group D – 26.01.24 until 28.01.24, Meißen) and also *"from a scientific perspective to understand what is possible in the real world, where we also need and can go with our students, to have a justification for completely different teaching concepts or ideas about teaching and training. In addition, I found them very exciting from a content perspective, due to my combination of subjects"* (Group B – 24.01.24 until 26.01.24, Meißen). Furthermore, the guest lectures also helped or can enable to show students the opportunities and risks of AI and also to enthuse them for these topics or corresponding occupations.

Format – Content: Neutral: Technical content wasn't as interesting as application content, unclear work instructions not problematic, Abalone: good, but questions remain for Orange3

It was noted that the technical/theoretical content was not as interesting as the application/practical content. While it was noticed that task assignments were unclear, to some extent, this was seen as unproblematic: *"Naturally, I can view this from the perspective of a teacher in a completely different way. But in this case, we are not children, but adults, and we can work with task assignments that are not perfectly*

thought out" (Group B - 26.01.24 until 28.01.24, Meißen). The Abalone exercise was assessed as a good start, but questions remained unanswered, especially regarding Orange3.

Work instructions were described by other interviewees as too unclear: *"They were not good, confusing, misguided"* (Group B - 26.01.24 until 28.01.24, Meißen). Overall, there was insufficient differentiation in terms of training content being specific to the field. Transferability was questioned for training's content that referred for example to archaeology and natural sciences (e.g., guest input / lectures and Abalone). Also, there often wasn't enough time for discussions or practical exercises. One participant felt that practical exercises were unnecessary.

Format – Content: Negative: Work instructions unclear, confusing, not (very) good etc., lack of specific subject didactic reference, subject-specific content not transferrable for other subjects, too little time for practical exercises and discussions, Practical tasks superfluous

Additionally, it was noted that although a lot of knowledge was gained through the training, there is a feeling that practical work for the teacher begins only after the workshop. It was also criticized that the term AI was not sufficiently defined: *"With AI, it remained a bit open until the end, […]. Perhaps it would have been useful to provide a working definition, perhaps it was intentional that it remains a bit open for us. I like to have something to hold on to. Maybe it is not yet scientifically secured. It is not new, but currently it is still not foreseeable what exactly it is and how it can be limited"* (Group D - 26.01.24 until 28.01.24, Meißen). Regarding the definition of terms, it was expressed that they came too late in the training process.

Format – Content: Negative: AI not so clearly defined, Theoretical concepts / terms / definitions introduced too late

When criticism was addressed regarding specific exercises, it was noted that the development of concrete proposals within the AI & Data segment on Sunday was insufficient, or there were unnecessary

Format – Content: Negative: AI & Data with problems, Teaching Workshop: not long enough, AI chatbots in schools, Grimm's New Fairy Tales with unclear work instructions,

parts included: *"For the competency requirements for teachers, personally, I was not interested in how someone arrives at them. It would have been sufficient for me to see the competency model once and then engage in discussion about it [...]. Comparing different models is very important for someone researching didactics but absolutely irrelevant for us as users. I found that very unfortunate because it took up a lot of time that was sorely needed elsewhere"* (Group C - 26.01.24 until 28.01.24, Meißen). Similarly, for the subsequent Teaching Workshop, participants wished for more time. Concerning the exercise on AI chatbots in schools, it was criticized that the results were not clear or meaningful. Regarding Grimm's New Tales, inaccurate task assignments were criticized, partly leading to a *"[...] complete blackout and not knowing exactly what to do"* (Group A - 26.01.24 until 28.01.24, Meißen). The Miro board tool was deemed insufficient for ensuring results, as it was not consistently used by all participants and was often switched to other (hybrid) forms by the workshop facilitators.

Miro Board not optimal / very good

Regarding the guest lectures or presentations, it was critically noted that in some cases, not everything could be presented as intended by the speaker, and there were also no practical connections/exercises directly following these lectures. The sequence of two presentations as well as the city tour was also perceived as too much consecutive theoretical input.

Format – Content: Negative: Guest input (presentation/lecture) with problems and no practical connection to guest input

The training plan was generally considered well designed in theory, *"[...] but the implementation was obviously hardly possible because it was too slow"* (Group C - 26.01.24 until 28.01.24, Meißen). The time management during the training was therefore classified as difficult, often requiring cuts, haste, or short timeframes at certain points.

Format – Length: Negative: Training too short, design/plan too full for available time, time management difficult, had to rush through content

Therefore, there was sometimes the feeling of undergoing the training under time pressure.

Other factors: Positive: Training design / plan (very) good, exchange with other teachers (very) good / interesting

Positive feedback was given on other factors, as just mentioned, praising the fundamental design/plan of the training. It was also emphasized that the collegial exchange during and around the training was highly appreciated, for example: *"For me, as someone who had no prior knowledge, the best part was actually the many conversations during the breaks with colleagues who already know a lot. That's how I learned a great deal, probably the most"* (Group B - 26.01.24 until 28.01.24, Meißen).

Ideas for further training: Training-specific (Overall training) – Clearer design / plan

Regarding areas for improvement, even though, as just illustrated, the training plan was praised, it was recommended to *"[...] establish an even clearer roadmap, even if it may not fully encompass the current scientific discourse, but at least one can work with it concretely"* (Group C - 26.01.24 until 28.01.24, Meißen).

Ideas for further training: More support for / at schools regarding joint integration. Bring in experts.

It is desired that, apart from the educational federalism in Germany, there would be more consensus and support among/at schools regarding the integration of (new) topics. Additionally, it would help *"if we really want to prepare students for life in the future work and research, [...] to bring in real experts to the school. These could help teachers and students to truly understand this process"* (Group A - 26.01.24 until 28.01.24, Meißen). This would counteract the process whereby every teacher / each school would have to create their own materials and be left to their own devices.

Ideas for further training: More practical exercises, content re-

In terms of content, some participants expressed a desire for more practical exercises, greater relevance to application, and more

on didactics, such as *"[...] how to directly implement AI in schools, [...]. That's where it got really exciting for me [in the training]. It was noticeable how suddenly more people engaged, got into conversation, and a lot more discussion arose compared to just focusing on the technical aspects"* (Group A - 26.01.24 until 28.01.24, Meißen). There was also a wish for *"more of a laboratory atmosphere, where one can experiment. I noticed we started googling for suitable programs. It would be good if, once you've found something, you could go into it and see what it can do. Of course, now that's something you have to do at home. It would be nice to have a bit more time for that [in the training], I think"* (Group D - 26.01.24 until 28.01.24, Meißen). Additionally, there was a desire expressed for content regarding differentiability or adaptability to various students' grade levels. Furthermore, the idea was raised to expand subject-specific discussions and to receive input from experts or scientific insights on this.

garding classroom application level and differentiation; align training more towards laboratory (more trying out on own), more discipline-specific/subj input/discussions with colleagues

Clearer working definitions would have helped: *"The basic terms were not specified. We didn't know which terms to work with because they weren't defined yet. Creating a working definition for ourselves, providing more assurance phases, and perhaps working more based on interests"* (Group B - 26.01.24 until 28.01.24, Meißen). This was particularly related to AI. During the working phases, it was suggested to better utilize Padlets instead of relying on Miro boards working with Padlets[5], which *can "[...] really break it [content] down so that you really have something and not just a framework where half is missing"* or TaskCards[6], *"[...] where there is better visibility and easier accessibility"* (Group D - 26.01.24 until 28.01.24, Meißen). The work instructions

Ideas for further training: Create definitions of terms/concepts/definitions in training, provide working definition of AI, use Padlets or TaskCards instead of Miro Board, clearer work instructions (especially for students)

[5] See https://de.padlet.com (no English version available)

[6] See https://www.taskcards.de (no English version available)

should be designed more clearly, especially with the perspective or the requirement if they were to be adopted for students in the classroom.

Ideas for further training: Clearer work instructions for Grimm's New Fairy Tales and Abalone, Grimm's New Fairy Tales (more time for assessment), Abalone (cover more in depth, especially Orange3), Presentations (as early as possible and handouts)

For clearer work instructions, it was also advocated elsewhere, particularly in the case of the exercises Abalone and Grimm's New Fairy Tales. Additionally, for the latter exercise, more time for assessment would have been desirable. For the Abalone exercise, a more in-depth exploration should have been done, especially regarding Orange3, as some questions remained unanswered afterwards. Presentations should ideally have a stronger practical relevance to classroom teaching. Any presentations that are theoretical in nature should be held early in the workshop to establish a working basis, and there should also be accompanying handouts.

Ideas for further training: Wish for another weekend workshop (e.g., for one more day)

Regarding length of the training, desires were expressed that justify a further weekend workshop, such as *"[...] to participate in such a weekend again, which was very enjoyable, if afterwards one could receive even more concrete things that could be taken as an add-on into one's own teaching"* (Group C - 26.01.24 until 28.01.24, Meißen). Additionally, it was also frequently expressed to want to spend more time on training or specifically to extend the weekend workshop, which lasted two and a half days, for example, by adding another day. One-day workshops, on the other hand, are insufficient for taking away something tangible and applicable.

Ideas for further training: Outsourcing of content (such as lectures) to save time, more time for exchange with colleagues

Further time-related suggestions include the possibility of outsourcing theoretical inputs, meaning providing participants with access to them before the training, such as in the form of videos. Ad-

ditionally, there was a desire for more time to be allocated for collegial exchange. Furthermore, activities related to subject-specific or school-specific application (like AI and Data and the Teaching Workshop) should take up more time in the training.

and for specific (applicational) exercises / content

Linked to the desire for an extended weekend workshop was the demand for more concrete teaching materials. Also, scientifically grounded, proven materials containing concrete steps were brought into play: *"For example, even at the university, if experts are not invited into schools, things could be further developed, including entire series, which are then evaluated and handed over to teachers. Following the motto: 'You don't have to develop it completely on your own. Here is something I have tested and works well. Take this.' [...] I would wish for more scientific background to be provided"* (Group A - 26.01.24 until 28.01.24, Meißen). Teachers could also give feedback for this kind of material. This would contribute to relieving the burden on teachers. Additionally, a prompt catalogue was requested within the exercise Grimm's New Fairy Tales, to make it clear what needs to be provided to obtain more concrete results. For the presentations, there was a request for handouts.

Ideas for further training: materials: more concrete for own teaching, well-founded material, concrete implementation (ready-to-use) plans, Grimm´s New Fairy Tales: catalogue for prompts, presentations (e.g., guest lectures): handouts

Summary:

The overall feedback across CS, STEAM, and primary education training sessions reveals a consistent theme: while the content and structure of the training are well-received, there is a clear call from all participant groups for longer durations or series of trainings and deeper material engagement. This would better accommodate the needs for detailed understanding and practical application necessary for teaching AI and DL effectively.

For CS teachers, the sessions ranged from 90 minutes to 7 hours and received generally positive feedback regarding the introduction to AI and DL, with the material mostly aligning well with participants' pre-existing knowledge. However, responses varied concerning the tools used, such as Orange3, and the length of the trainings. Some participants felt the trainings were too short to adequately cover all the material, while others found them overly lengthy. Suggestions for future training emphasized the need for more in-depth content and longer, more detailed sessions. CS training in Germany extended over 2.5 days and was noted for its thorough engagement with the material, providing a robust foundation in AI and DL, although some participants suggested an even more extended period (a series of workshops) could be beneficial for deeper exploration and understanding.

STEAM teachers experienced sessions lasting between 1.5 to 4 hours. Like the CS group, they appreciated the training structure and the integration of practical and theoretical components. Nonetheless, they echoed the need for adjustments in the lengths to better accommodate deep dives into specific topics and specific applications for STEAM subjects that they were lacking. The feedback suggested a preference for spreading the trainings over weeks or months. In the third round of STEAM training in Germany, spanning 2.5 days, participants also reported a positive reception of the content that combined AI and DL, which was considered to align well with their prior knowledge. This longer format allowed for a more in-depth exploration of the subjects, appreciated by the attendees. Similar to other groups, however, there were suggestions for even more extended sessions. Participants valued the comprehensive nature of the training but expressed a desire for further depth and more practical applications that could be directly transferred to classroom settings. The longer duration was seen as beneficial for delving deeper into complex topics and providing a more sustained learning experience, which was particularly useful for integrating the new knowledge into educational practices effectively.

Primary educators, who attended 3-4 hour sessions in the third round, generally found the training well-aligned with their knowledge levels and the length appropriate. They expressed a desire for longer and more frequent training opportunities to better integrate AI and DL into their curricula. Practical tasks were well-received, but there was a call for more interactive and hands-on approaches.

3.3 Teachability and difficulties of teaching DL and AI

CS in all rounds, length 90 minutes – 7 hours:

In all rounds of CS teacher training, the urgency to integrate DL and AI into the CS framework curriculum was consistently emphasized by teachers from various countries. Teachers also advocated for the incorporation of DL and AI into non-CS subjects and proposed making CS a mandatory subject in secondary education across Germany and Austria, with some in Lithuania suggesting starting CS education at the primary level. Teachers discussed the difficulty of moving from basic theoretical knowledge to applying DL and AI in more significant, contextual settings in the classroom. The gap between teaching foundational concepts and applying them in complex tasks was particularly challenging. Teachers suggested that starting with simple, unplugged exercises to explain basic principles before advancing to larger models could help bridge this gap. The integration of DL and AI is not only a technical issue but also an ethical one. Teachers need additional contextual knowledge to address ethics.

STEAM 2-3 rounds, length 1.5-4 hours:

From the second round in Lithuania and Austria, it was noted that the training content was adaptable for subjects like biology and mathematics, with modifications suggested to make the content accessible for varying student age groups. The training was seen as feasible for younger students through playful approaches and could be scaled up in complexity for older students. Teachers showed a keen interest in integrating these technologies into the classroom, driven by the motivation to update and make their subjects more engaging and relevant.

In the third round, the necessity to adapt the training and educational materials to different educational levels and subjects was emphasized, with a focus on making them more accessible and relevant. Practical applications were favoured, and the ability to implement simpler exercises in lower grades and more complex tasks in higher grades was considered crucial for effective education in DL and AI. For both rounds, more practical applications for teaching DL and AI into STEAM subjects are needed.

Primary 3rd round, length 3-4 hours:

Teaching DL and AI in primary school education poses challenges: For effective instruction, teachers need knowledge, and external support from outside the school could be helpful. The teachability varies based on topics and grade level. Emphasizing playfulness, such as through puzzle tasks and practical activities like using Bebras cards or tasks from the ViLLE platform, is essential. DL is considered teachable for lower grade levels in general. While AI's feasibility in primary education is also acknowledged, creating tasks may be challenging. Possible approaches for teaching AI include working with images, addressing fake news, and emphasizing pattern recognition. A connection between DL and mathematics is highlighted, aligning with competencies taught in this subject.

CS Germany 1st round, length 2.5 days:

As a prerequisite for the teachability of DL and AI, it was mentioned that students would first need basic knowledge to be able to handle and learn more about these topics.

Teachability: DL & AI – Dependent on the depth of the treatment, Jupyter Notebooks: Feasible (e.g., data cleansing)

It could be difficult for the students to evaluate tasks such as surveys. Such tasks are not easily conveyable.

Teachability: DL – Challenging / difficult

As seen in the exercises during the training, it is possible to go through processes related to AI entirely with the students. Therefore, teaching AI, even if the content is further deepened, is feasible in classroom teaching. However, the content would also need to be tailored to the students.

Teachability: AI – Feasible, Dependent on students or their characteristics (e.g., age)

Students would quickly become interested in things they find on the internet, such as ChatGPT. However, as a teacher, there would be the challenge of explaining the complex functionality of this tool. Mathematical derivations in combination with neuron functioning would be teachable at the upper secondary level. Similarly, exercises or games based on decision trees would be feasible in the classroom, which would be suitable for younger students.

Teachability: AI – Students are interested in ChatGPT, ChatGPT: complex, mathematical derivation and neuron functioning for upper secondary level, decision trees / rule setting: feasible for younger students

If teaching is implemented well and the teacher can inspire the students, there is a high teachability, even for young students. Specifically, the programming language Scratch or programming in general was considered suitable for younger students. For the conveyance of content, a way must be found so that students are neither under nor over challenged. It can be difficult or time-consuming to convey the content, but fundamentally, almost any material can be taught.

Teachability: Unspecified – Feasible, Scratch & programming: feasible for younger students, neither under- nor overburden students, challenging / difficult

As mentioned just before, the conveyance of content also depends on the teacher – their enthusiasm and the topic implementation play a role. Additionally, the teachability depends on the depth of treatment: *"You always just have to look at the depth, in which class. I don't see any problem with that at all, because we also teach them computer hardware or programming, and that doesn't mean they have to be*

Teachability: Unspecified – Dependent on teacher, how teaching is implemented, depth of the treatment and on students' grade levels

able to program some high-end super software, but it's about the basics, the understanding" (Participant C - 25.11.22 until 27.11.22, Heidelberg).

The lower the grade level, the more comprehensible the content must be conveyed, because if it becomes too complex, the motivation of the students can decrease.

Teachability: Unspecified – The lower the grade level, the more understandable the teaching; when too complex, motivation can diminish

Regarding ideas on how to integrate the topics into the classroom, with the Donkey Car Pilot[7], you can bring these topics into the classroom: build vehicles, conduct training runs via the app, etc., and thus achieve relatively quick success.

Ideas how DL & AI could be taught: Donkey Car Pilot

There are some data for school use that are already prepared and available, such as historical climate data. Such or other data could be integrated very well into various school subjects: *"Yes, science or our modern world, we are sitting on enormous amounts of data nowadays, and we just need to prepare them in a way that they are accessible even for ten, eleven, twelve-year-olds"* (Participant B - 25.11.22 until 27.11.22, Heidelberg).

Ideas how DL could be taught: Much (other) data sources for school, working with data from climate institutes, much data out there for integration in many subjects

One way to integrate DL into the classroom is to take photos via GPS positioning, load them into a map, and evaluate the attractiveness of these locations (for example, using green, yellow, or red dots) based on data such as climate data. Another idea is for the students to collect data related to their school (e.g., quality of bicycle paths) and then enter it into a web-based application, for example. Furthermore, students could clean data using Jupyter Notebooks.

Ideas how DL could be taught: Environmental and spatial visualisation (e.g., per photos, GPS), collecting data in one's own school environment, Jupyter Notebook: Data cleansing

[7] See https://docs.donkeycar.com/guide/train_autopilot

With the QuickDraw[8], where users are given prompts to draw and the tool then recognises whether the prompt has been drawn, there is a playful approach to integrating AI into the classroom. For example, it could be incorporated into language classes, where the class has to guess vocabulary based on the drawings. Additionally, the teacher can share background information about AI with the students. Another possible approach to AI is Teachable Machines, for instance, embedding tasks related to relevant social contexts, such as recognising whether people are wearing masks (Corona epidemic) in photos.

Ideas how AI could be taught: Tool QuickDraw (e.g., work with vocabulary in foreign languages), Teachable Machine: ethical, social issues (e.g., corona mask)

An idea with real-world relevance for students is programming chatbots with Scratch: for example, creating a chatbot similar to those encountered by users on Amazon. Furthermore, Jupyter Notebooks can be used to address a variety of general problems.

Ideas how AI could be taught: Programming a chatbot (with Scratch), Jupyter Notebook: character recognition

The relevance of the topics DL and AI is assessed in such a way that it is important to teach these subjects because they are very present in everyday life.

Relevance of DL & AI: General statement - (very) high / Important

When it comes to discussing the linkages of DL to individual subjects, thematic connections to the subjects of biology, physics (STEAM), and geography were identified in the interviews, such as: *"Geography is naturally very media-intensive and data-intensive anyway. I would count things like collecting our own data in the vicinity of the school on the quality of bicycle paths, for example, and then entering that into special web-based mapping applications either on mobile devices or using various cool tools [...]"* (Participant A - 25.11.22 until 27.11.22, Heidelberg).

Linkage between subject to DL: Biology, physics (STEAM), geography: yes - data-intensive

[8] See https://quickdraw.withgoogle.com/?locale=en_US

In programming languages like Scratch, one can rebuild chatbots, for example, in a foreign language subject. Another example is the previously mentioned QuickDraw, which can be integrated with language exercises, such as vocabulary learning.

Linkage between subject to AI: Foreign Languages: yes - Text-based AI / tools

CS Germany 3rd round, length 2.5 days:

As a prerequisite for the teachability of DL and AI, it was mentioned that students would first need basic knowledge to be able to handle and learn more about these topics.

Teachability: DL & AI: Prerequisites - Imparting basic skills / competences

The students are easier to motivate by AI compared to DL*: "We have databases or data and their structuring as a content field in [federal state], but I find it difficult to motivate why data modeling is important. Through the topic of AI, one can effectively demonstrate that modeling car classes is not done just for fun. The students do not understand the relevance of it, but here the effects are clearly visible and can be made tangible"* (Group A - 24.11.23 until 26.11.23, Zeitz).

Teachability: DL & AI – Students find it easier to be motivated by AI than by DL

In relation to exercises from the training, the exercise food Inspection (bias/image recognition) would be applicable in class.

Teachability: DL & AI – Food inspection (algorithmic bias / image recognition)

Decision trees could be implemented as part of the exercise Good Monkey – Bad Monkey and the use of the tool Orange3.

Ideas how DL & AI could be taught: Good Monkey – Bad Monkey and decision trees together with Orange3

In relation to DL, using tools like Orange3 as an example that is applied in training, it was mentioned as a prerequisite for teachability that exercises must be constructed in such a way that students

Teachability: DL: Prerequisites - Orange3: set up in a way that they can learn something

also understand why it works and can then apply this knowledge later on.

Teachability of DL: Dependent on good examples, that create motivation; DL und data concepts through databases; feasible for lower grade levels

Teachability in the field of DL thus depends on finding good examples/tasks that engage the students, as otherwise these topics can quickly become dry. DL as a fundamental concept of data could be effectively conveyed through the topic complex of databases. Certain content in the field of DL would be teachable for eighth-grade students.

Teachability of DL: Orange3 for higher grade levels; pizza project (data workshop) and Generative Cocktail Producer: May be feasible for upper secondary level, but not lower secondary level

Regarding the teachability related to training content from the weekend workshop, it would be conceivable to use Orange3 with eleventh graders in the subject of mathematics. The Pizza project (data workshop) would be conceivable for the upper secondary level but not for the lower secondary level. The same applies to the exercise Generative Cocktail Producer.

Ideas how DL could be taught: Orange3: work on functions unplugged, then introduce plugged; usage of InstaHub to motivate students to work with data

In the interviews, some ideas were brought up on how to teach DL. The program Orange3, for example, could be introduced first through an unplugged approach: *"I would start by doing it manually so they understand what's happening. For example, you draw the points and make a line through them with the triangle ruler. 'Now we have a tool that can do it better!' Then I would move on to the tool so they understand what's happening. Otherwise, it's a black box where I have a group, data, and a line [in Orange3], and suddenly a finished result comes out predicting something. It's an abstract black box, and in the worst case scenario, students learn that they have a program that does it. I want them to understand*

what's happening" (Group A - 24.11.23 until 26.11.23, Zeitz). With the social network online tool Instahub[9], students could be motivated to work with data and analyse it.

Data security could also be addressed outside of CS classes, for example in Ethics subjects or in the class council. Since data plays a role in many subjects, it can be referenced and worked with accordingly there, for example climate or measurement data in physics (for example, via sensors).

Ideas how DL could be taught: Data security can be taught in Ethics or class council, data competences in many subjects (e.g., climate data in physics)

Regarding the teachability of AI, ChatGPT could be utilised, for example, as a supportive tool in programming tasks within the CS subject for grades nine to ten. Some of the exercises from the AI training would be suitable for students.

Teachability of AI: ChatGPT for lower grade levels , feasible

As just mentioned, ChatGPT can be used in the context of programming: *"But one must always point out that it needs to be adapted because we use an online programming environment where individual methods are often different. As support, it [ChatGPT] is good, but it does not replace understanding"* (Group A - 24.11.23 until 26.11.23, Zeitz).

Ideas how AI could be taught: ChatGPT for programming

After the training, AI is defined in relation to the use of linear regression for determining the pizza price as follows: *"[...] you input some data and something reasonably sensible comes out"* (Group A - 24.11.23 until 26.11.23, Zeitz).

Definition of AI

The prerequisite mentioned is that for teaching content, it must also be presented and designed in an interesting manner.

Teachability (unspecified): Prerequisites - Content must be explained / designed interestingly

[9] See https://instahub.org (no English version available)

During project weeks, one can work on topics and also break them down for different grade levels, making it interesting for each respective grade level.

Ideas how topics could be taught (unspecified): In project weeks, topics can be worked on (for different grade levels)

Concerning relevance, DL and AI are important topics for teaching and just as relevant for integration into the classroom as all other topics that are equally significant for students.

Relevance of DL & AI: (Very) High / important

Furthermore, it is assessed that *"[...] especially with the increasing amount of data, which is growing and which we can and must process ever faster, this data competence becomes increasingly relevant for students, because only by analysing such data can they learn to properly assess certain results they receive, for example, on Instagram or TikTok"* (Group B - 24.11.23 until 26.11.23, Zeitz). This way, students can be educated towards independence. For another respondent, programming algorithms is the most relevant aspect in the subject of CS, followed by DL.

Relevance of DL: data Competences getting more important for students;
in CS, programming (algorithms) are the most relevant before DL

The handling of AI is a sociopolitical and social science issue, therefore, integrating it into education is important.

Relevance of AI: (Very) important (e.g., because social, occupational factors)

Relevance and thus linkage of DL and AI are seen across CS and extend to other subjects; for example, DL and AI *are "[...] ultimately also part of our lives, and I believe that other subjects such as German or religion or others can and should address the topic as well"* (Group B - 24.11.23 until 26.11.23, Zeitz). There are applications and examples from all areas, and accordingly, impacts will unfold in many contexts.

Linkage between subject to DL & AI: Non-CS-subjects – yes, linkages / examples can come from all subjects / areas

Therefore, interdisciplinary work in subjects other than just CS should be done on this topic.

In mathematics, fundamentals are taught that can be utilised within the scope of AI and DL (e.g., modeling).

Linkage between subject to DL & AI: Mathematics

Data plays a role in environmental protection. This topic is covered in several subjects, such as geography or physics, through (climate) data measurements that can be obtained using sensors. In the subject of physics, working with data (such as data preparation) is relevant. Regarding the STEAM subject of mathematics, a connection to DL is seen in regression analysis, but also in statistics. The latter also occurs in the natural sciences.

Linkage between subject to DL: Non-CS-subjects – data; STEAM (Mathematics, nature sciences, physics) – statistics; in physics, working with data ; Mathematics – regression analysis, data / statistics

Aspects like data protection do not have to be addressed in the CS subject, but can be covered in ethics classes, for example.

Linkage between subject to DL: Non-CS-subjects – Ethics - data security

Regarding AI, it was said, among other things, *"the content context is in other subjects because AI is not an end in itself, but it serves an application, and this application is within the realm of the geography teacher or the German teacher"* (Group B - 24.11.23 until 26.11.23, Zeitz). Therefore, AI does not necessarily have to be taught exclusively in CS but can also be integrated into other subjects, as statistical tasks are also associated with it. Furthermore, it should be considered whether AI could be integrated *"[...] just as well as societal aspects into other subjects, where it is viewed from different perspectives. We need to address it in school in general, in various areas. In this, computer science may have a small role, where students actually learn what is behind it. However, much more relevant is the fundamental societal approach to it"* (Group C - 24.11.23 until 26.11.23, Zeitz).

Linkage between subject to AI: Non-CS-subjects – yes, can be brought as societal aspects in subjects

The topic of AI can be incorporated into mathematics lessons within the framework of statistics.

Linkage between subject to AI: Mathematics – connection to data / statistics

STEAM Germany 3rd round, length 2.5 days:

Modelling experiments like the Abalone exercise could be considered for advanced courses in both biology and chemistry at the upper secondary level. Unplugged exercises (such as Grimm's New Fairy Tales) would be suitable for teaching as they are independent of technical malfunctions or difficulties.

Teachability of DL & AI: Modelling experiments like Abalone for upper secondary level and advanced courses (biology and chemistry), unplugged games are technically independent

The Abalone exercise could be integrated into the subjects of biology and chemistry within the context of ecological topics at the lower secondary level.

Ideas how DL & AI could be taught: Abalone for embedding in ecological themes

For the teachability in the field of DL, the dependence on time was emphasized

Teachability of DL: Dependent on the available time in class

Topics like AI are more exciting for students than mathematics. In relation to the lesson, the students should be able to evaluate output: *"I always question whether the students' subject competence needs to be even higher. I can input any prompts into ChatGPT, but I must be able to assess whether the result is usable at all. I think that students need to be sensitised. They must ask quite specific questions and they need expertise to evaluate the answer"* (Group D - 26.01.24 until 28.01.24, Meißen). Furthermore, it should be made clear to students what AI can achieve for a specific case or exercise and where its limitations lie.

Teachability of AI: Students are more interested in AI than in mathematics, lack of required knowledge among students, make clear what AI can't do, where the limits are

When teaching AI, the aspect of fun could also be emphasized, and perhaps, *"[...]AI can be just funny for once. You train it and fundamentally learn what AI actually does. You do this exercise with the monkeys [Good Monkey-Bad Monkey exercise] beforehand, deconstructing sentences, and then finish with GenAI. You just have a bit of fun. You can choose what to train it with, and then it's tested against each other. I could imagine something like that immediately"* (Group D - 26.01.24 until 28.01.24, Meißen). Furthermore, for future teaching, it is important to convey differences between AI and non-AI-based applications/tools, such as between Chat GPT and querying in search engines like Google.

Ideas how AI could be taught: Fun exercises with AI / Generative AI, comparing tools / applications that use AI and don't

With AI, experiment protocols could also be created. The experimental generation of fake news (as text, as an image, or in combination) can help to reflect on the functioning of AI. On the other hand, it might be *"[...] a bit more fun to deconstruct AI-generated texts, to see if it fits or if it might need optimisation. [...] I always feel that the correction is a bit nicer pedagogically when it concerns something artificially generated, rather than from works of other people you know"* (Group B - 26.01.24 until 28.01.24, Meißen).

Ideas how AI could be taught: Writing experiment protocols / procedures, produce fake news (as text, image etc.) as trial, deconstruct texts with AI instead of creating texts

The workshop has created awareness that it is also important to convey and understand what AI cannot do yet.

Definition of AI

Even if the students (or the teacher) do not yet know something, the students are trusted to be able to work out the knowledge themselves.

Teachability (unspecified): Students can gain knowledge by themselves

There shouldn't be too many fears among teachers that they can't teach something about these new topics: *"Essentially, you have to send the students on a research journey themselves. If we forget what the next*

Send students on "research trip" (they'll explore themselves)

step was, they'll [students] figure it out somehow. You don't necessarily have to be an instructor standing up front and saying, '[Do it] like this and this and this'" (Group D - 26.01.24 until 28.01.24, Meißen).

Relevance of AI: Biology - AI has relevance

For the subject of biology, AI is considered relevant, *"[...] because we also conduct experiments there. I mean, even science works with AI today. Why not start with it in school already and be progressive and do something that is already further in science. You can then introduce them to it earlier as well"* (Group A - 26.01.24 until 28.01.24, Meißen).

Relevance of AI: Humanities - No (immanent) need to use AI

For the subject of humanities, on the other hand, there isn't necessarily seen a need to use AI for teaching at the lower secondary level.

Linkage between subject to Unspecified: (Foreign) Languages: yes - Text based work which can be individualised to students

In the subject of German or German as foreign language, strong adaptations can be made to meet the needs of the students, and many interactive exercises can be conducted – although the interviewee does not specify exactly what these exercises consist of.

Linkage between subject to AI & Unspecified: Foreign Languages: yes - Social Media platforms

In English, there is the topic of media society*: "There, one can definitely connect. They [students] constantly use social media platforms without really knowing or understanding the algorithms behind them, at least in the basic principles. This is important both for skills, that is, knowledge about AI and how these algorithms work"* (Group A - 26.01.24 until 28.01.24, Meißen).

Linkage between subject to AI: Non-CS-subjects: yes - For things with potential / clear solution path (e.g., design)

In general, for non-CS-subjects, it is said that when it comes to certain solutions such as in the topics of design, fashion, and engineering, where text areas or technical drawings are involved, AI could be used.

On the other hand, it could be difficult for an AI if solution paths are very branched out.

Linkage between subject to AI: Non-CS-subjects: no - For branched solutions / problems, AI is unsuitable

There is the opinion that integrating topics of DL and AI in the subject of music is difficult: *"I have received a few ideas, but I still need to familiarise myself more with the specific programs available or engage more with them. I need to see what is possible. What is the added value for the students? Not that I enter something, a product comes out, and that's it"* (Group D - 26.01.24 until 28.01.24, Meißen). Therefore, while the applicability is seen in the subject of CS, it remains unclear for the subject of music.

Linkage between subject to DL & AI: Music: no - unclear where / if it can be applied

On the one hand, it would also be possible to address, among other things, *"[...] that AI is used to generate music independently of humans, to then analyse it, and potentially establish connections to existing compositions written by humans, as well as addressing copyright issues. There are many possibilities"* (Group B - 26.01.24 until 28.01.24, Meißen).

Linkage between subject to AI: Music: yes - generation of music

On the other hand, in the subject music, the connection to AI can be established by discussing how songs are suggested.

Linkage between subject to Unspecified: Music: yes - How songs are suggested

For creating images using LLMs, no potential application is seen in the subject of art, as the results or the creation of images using prompts would lead to unpredictable outcomes.

Linkage between subject to AI: Art: no - Creating images with LLM unpredictable or unachievable results

In biology, there are linkages from DL and AI to behaviour predictions, genetics, and modelling experiments such as the Abalone exercise.

Linkage between subject to DL & AI: STEAM - Biology: yes - Behaviour predictions, genetics, modelling experiments (like Abalone)

In mathematics, on the one hand, a connection to application in linear regression was identified. On the other hand, it's about optimizing and modelling, although it was not further elaborated on what this exactly entails.

Linkage between subject to DL & AI: STEAM – Mathematics: yes - Usable for topics connected to regression, optimising and modelling play a central role

The same interview mentioned that in Physics optimising and modelling play also a central role.

Linkage between subject to DL & AI: STEAM – Physics: yes - Optimising and modelling play a central role

In chemistry, the interviewee could imagine fewer applications for collecting data than in biology.

Linkage between subject to DL: STEAM – Chemistry: Application areas for working with data harder to find than in biology

Summary

In the discussions on the teachability of DL and AI across CS, STEAM, and primary education, several themes emerged. Teachers consistently recognized the importance of integrating DL and AI into teaching. Challenges highlighted included bridging the gap between theoretical concepts and their practical applications in the classroom. The feedback suggested starting with simple, unplugged exercises to introduce basic principles, and gradually moving to more complex models. For younger students in STEAM and primary settings, the approach recommended was more playful and interactive, adapting the complexity of content to be age-appropriate. Teachers in STEAM and primary emphasized the need for application linkages of DL and AI to their subjects.

3.4 Integration experience before the training and ability to integrate D and AI after the training

CS in all rounds, length 90 minutes – 7 hours:

Participants from all three countries noted a gap between the training received and its practical application in the classroom. Teachers felt better prepared after the training but recognized that effectively incorporating DL and AI into their teaching would require further engagement with the material. A single 7-hour session was deemed insufficient for teachers, particularly those who had not previously taught these subjects. There was a strong call for more concrete course plans and materials, specifically tailored to different grade levels, to facilitate easier integration into teaching.

STEAM 2-3 rounds, length 1.5-4 hours:

AI topics have been integrated into classroom settings with mixed results. Positive experiences were reported in an English class using ChatGPT, which sparked student interest. Conversely, another instance involved using chatbots, but it did not lead to significant subject-wise learning or analysis by the students. AI has been used to demonstrate mathematical models in various educational settings, including gymnasiums, where limitations were noted in deepening the exploration of the subject. Barriers to the broader integration of AI include legal and safety concerns, particularly for younger students in primary schools, and the complex nature of AI, which some find daunting. Despite these challenges, there is an expressed interest in future integration of these technologies. Effective implementation of DL and AI in teaching is seen as feasible but requires significantly more support, materials, and detailed guidance. The current training only provided basic knowledge, underscoring the need for ongoing learning opportunities to enable teachers to effectively convey these complex topics in the classroom.

Primary 3rd round, length 3-4 hours:

Post-training, the participants recognized a significant gap in their ability to apply what they learned effectively in their teaching, pointing to a shortage of in-depth factual and didactic knowledge and a lack of external expertise. While the training provided a foundational understanding, it was insufficient for immediate application in the classroom without additional support and resources. Some participants felt ready to immediately integrate specific concepts like decision trees, whereas others anticipated gradually introducing more interactive elements such as card games in future lessons.

CS Germany 1st round, length 2.5 days:

Participant A has already designed citizen participation projects, although it was not clear to what extent students (in the subject of geography) were involved. Real pictures were taken, then positioned on a map via GPS data, and the attractiveness of the locations was evaluated based on various data.

Participant A: Integration: DL, subject unclear, frequency – more than once: Environmental and spatial visualisation (e.g., per photos, GPS)

In CS class, chatbots were programmed using Scratch. The interviewee attempted to establish connections to real contexts, such as chatbots for support found on sales platforms like Amazon, or for registration on other platforms, or for computer games. However, it remains unclear whether an AI was designed in the way it was applied in class; the chatbot was a *"[...] very basic approach, but it's quite nice that the students have the opportunity to essentially specify what exactly should it [chatbot] say in response to this or that question or statement, and to open up these different options"* (Participant A - 25.11.22 until 27.11.22, Heidelberg).

Integration: AI, subject CS, frequency – more than once: Programming a chatbot (with Scratch)

Teachable Machines were also implemented in CS: *"[...] the idea we then implemented was actually to train an AI application that recognizes when people wear their Corona [COVID-19] face mask incorrectly. [...]. At*

Integration: AI, subject CS, frequency – unclear: Teachable Machine, Ethics

the time, it was indeed a societal problem: people wearing masks below their noses or under their chins, etc. [...]. So, you can then train it [Teachable Machine] to recognise how the mask should properly fit over the nose and mouth. And then you can also demonstrate what happens when it hangs below the nose, under the chin, or when it's not worn at all" (Participant A - 25.11.22 until 27.11.22, Heidelberg). This task also raised ethical questions.

The Teachable Machine with the masks worked surprisingly well, even without needing to bring much knowledge about the algorithm's functioning.

Experiences: Positive (Teachable Machine) – Successful / good / very good

The next participant worked with data in the eighth grade, for instance, data was processed in a table, graphs of temperature trends were created, or box plots of rainfall data were discussed.

Participant B: Integration: DL, subject unclear, frequency – unclear: Working with data (e.g., analysing)

Additionally, in the subject of mathematics, the topic of linear classifiers is regularly integrated.

Integration: Unspecified, subject mathematics, frequency – more than once: Linear classifier

The participant (as CS teacher) collaborates with a biology teacher to cover the topic of neural networks in an elective course slot. For instance, a perceptron algorithm has already been addressed in this context.

Integration: AI, subject CS + biology, frequency – unclear: Neural networks

The lessons on neural networks have been successful so far.

Experiences: Positive (Neural networks) – Successful / good / very good

Participant C carried out projects at the upper secondary level, some of which involved elements of machine learning, such as a

Integration: AI, subject CS, frequency – unclear: Balcony solar

work on a balcony solar power plant, where codes were partially implemented by the students.

power plant (elements of AI)

Also, in the CS subject, Donkey Cars were implemented: *"There is a large community online, and students should replicate it as a project, so to speak. My goal was actually for students to understand less about what exactly happens in the background, but rather to implement projects that involve AI, and above all, for students to grasp this entire process, including data collection. [...] The network needs to be trained, it is played back, and then the vehicle can perform more or less well"* (Participant C - 25.11.22 until 27.11.22, Heidelberg).

Integration: DL & AI, subject CS, frequency – more than once: Donkey Car Pilot (can have elements of AI or/and DL)

The project works well, and successes could be quickly achieved in the past. The students also receive immediate feedback on whether something works or not, and they conduct data cleaning during the project.

Experiences (Donkey Car Pilot): Positive – Successful / good / very good; Students – (first) successes / results are achievable quickly

In Participant D's class, a Matchbox AI ("Streichholzschachtel-KI") was covered. Additionally, AI in general was a topic, but *"[...] only on an ethical level [...]. What is AI actually, and what is dangerous and what is good, and what are the opportunities, what are the risks?"* (Participant D - 25.11.22 until 27.11.22, Heidelberg).

Participant D: Integration: AI, subject unclear, frequency – once: Matchbox AI, Ethics

The students found the ethical aspect very interesting and engaging, and they were able to contribute a lot to the discussion in class.

Experiences: Students – Positive: Were interested (Ethics)

On the other hand, it was not possible to delve deeply enough into the (technical) topic matter: *"But it's no use without having a bit of depth and knowing technically what's going on. And when I assigned a presentation on neural networks, for example, it was always very, very superficial. So, I wouldn't even call that AI content. They looked at it and*

Experiences: Students – Negative: Not enough depth

thought, what is this?" (Participant D - 25.11.22 until 27.11.22, Heidelberg). In that sense, the treatment of the topic or the topic itself was difficult.

Participant E had already integrated content as well. There were multiple attempts made to teach classes on neural networks.

Participant E: Integration: AI, subject unclear, frequency – more than once: Neural network

But the teacher was somewhat uncertain whether the treatment of this topic in the class could be conducted or continued in that manner.

Experiences: Neutral: Teacher partially unsettled

Some of the participants would immediately feel capable of integrating content into their teaching after the training, for example: *"Well, personally, I already feel capable of implementing a certain introduction to AI in the classroom ad hoc, also based on the examples from the training. There were a few things there that you could definitely do with the students on an hourly basis in class, so you already have something"* (Participant C - 25.11.22 until 27.11.22, Heidelberg). Specifically, reference was also made to content from the exhibition "I am AI" that will be instantly integrated into the teaching by one of the interviewees.

Ability to integrate (after training): Instantly capable – AI: Instantly (as general statement), exhibits from the 'I am AI' exhibition / guided tour

An immediate integration of content on DL or AI from the training in the classroom would require: *"Well, I'm not the type of person who says okay, now I'm going to immediately use Jupyter Notebooks in my teaching or work on geography with Orange3 with my students, because that naturally requires an incredible amount of preparatory work and preparation time, which I may not even have"* (Participant A - 25.11.22 until 27.11.22, Heidelberg).

Ability to integrate (after training): Not (instantly) capable – DL & AI: Prerequisites: Continued engagement / preparation / time necessary

CS Germany 3rd round, length 2.5 days:

Regarding DL integration in the CS subject, participants report on the treatment of data competences in a ninth-grade class, where, for example, the data from a newspaper website were analysed (such as determining which company is behind it). Another participant emphasized in the ninth grade *"[...] concerning data, it is important to sensitise students to what data are requested by apps they install on their mobile phones, what data are collected by websites, and that it is not insignificant for them what kind of data are collected"* (Group A - 24.11.23 until 26.11.23, Zeitz).

Integration: DL, subject CS, frequency – unclear: Data competences

Databases are included in the curriculum for the ninth grade for natural sciences (as a part of STEAM subjects); however, further details on how the teaching was conducted are not provided.

Integration: DL, subject STEAM (natural sciences), frequency – unclear: Databases

Also, for subjects not specifically mentioned, databases are a topic in grades nine to ten. Here, there were mentions of data analysis that was also conducted, for example, so that students are able to form an opinion.

Integration: DL, subject unclear, frequency – unclear: Databases, working with data (e.g., analysing),

A participant once integrated AI into the CS subject, specifically addressing neural networks in a basic course.

Integration: AI, subject CS, frequency – once: Neural network

The same participant also discussed the ELIZA chat program as well as the Turing test with the students.

Integration: Unspecified, subject CS, frequency – once: Chat programs (ELIZA) in turing test

Now the reports on integration, which also included shared experiences, will follow.

A teacher from Group A tried out individual modules with students in preparation for AI teacher trainings. Further details were not provided. The modules could be easily implemented and they worked out fine.

Group A:
Integration: AI, subject unclear, frequency – more than once: Testing out modules with students for preparation regarding AI teacher trainings
Experiences: Positive: Content / teaching worked out fine

One teacher from Group B covers databases, including the work with Excel, which is regularly part of the participant's CS classes. Students generally handle this aspect well, and they can apply the knowledge gained. However, the acquisition of relevant datasets is always problematic.

Teacher 1 from group B:
Integration: DL, subject CS, frequency – more than once: Databases
Experiences: Positive: Successful / good / very good; Students – well understood and applied; Negative: data often not relevant enough for students

In a class with a higher grade level, another teacher from Group B provided students with an insight into AI: *"The students were very interested. However, I actually cannot draw on a really large wealth of experience. It will actually only become apparent with the implementation [meaning: how good it works] of the new framework curriculum in the 11th and then later in the 13th grade"* (Group B - 24.11.23 until 26.11.23, Zeitz).

Teacher 2 from group B:
Integration: AI, subject unclear, frequency – once: Small insight into AI
Experiences: Students - Positive: Were interested

In Group D, several participants shared experiences regarding the integration of DL and AI. The first participant introduces the topic of AI within the framework of mathematics lessons, specifically in the context of statistics; however, *"the students find it difficult, as with any other mathematical subject"* (Group D - 24.11.23 until 26.11.23, Zeitz).

Teacher 1 from group D:
Integration: AI, subject mathematics, frequency – more than once: AI in statistics
Experiences: Students - Negative: Some students did not come along / difficulties

Another participant mentioned neural networks in the class. The experiences shared were described as positive. Further details were not provided.

Teacher 2 from group D:
Integration: AI, subject CS, frequency – unclear: Neural network

Experiences: Positive: Successful / good / very good

In a double lesson of CS, the students played the game Good Monkey - Bad Monkey, with the following experiences: *"In the lesson, I felt that the students didn't completely understand how it relates to AI. I should have taken more time for that, not just a double lesson"* (Group D - 24.11.23 until 26.11.23, Zeitz). The topic of AI had to be incorporated separately because it is not included in the curriculum.

Teacher 3 from group D: Integration: AI, subject CS, frequency – once: Good monkey - Bad monkey Experiences: Negative: too little time; Students: did not fully understand it

In a double lesson of CS, the students played the game Good Monkey - Bad Monkey, with the following experiences: *"In the lesson, I felt that the students didn't completely understand how it relates to AI. I should have taken more time for that, not just a double lesson"* (Group D - 24.11.23 until 26.11.23, Zeitz). The topic of AI had to be incorporated separately because it is not included in the curriculum.

Teacher 3 from group D: Integration: AI, subject CS, frequency – once: Good monkey - Bad monkey Experiences: Negative: too little time; Students: did not fully understand it

Other participants expressed that they had not integrated the topics of DL and AI. One teacher mentioned not having taught these subjects yet in mathematics and physics.

No Integration – DL & AI: General statement

DL has not yet been integrated, as there have been a lack of datasets and ideas for its integration into classroom teaching.

No Integration – DL: Missing data(sets) / examples / ideas

AI has not yet been integrated, among other reasons, because the new framework curriculum in one federal state plans to do so only for the upcoming school year.

No Integration – AI: General statement, Missing data(sets) / examples / ideas

As prerequisites for the integration of the topics, it is stated that further engagement with the subjects is necessary to understand them theoretically, and additional knowledge would be required. Likewise, materials such as textbooks would be needed. For integration, content from current lessons would also need to be included: *"In principle, I already feel capable of adjusting my teaching accordingly [to integrate the topics of DL and AI]. I wouldn't have any major difficulties with that. My concern lies more in the fact that if I add something additional, I have to remove something else. For this, I need a thread of continuity for my teaching. I have to unravel it first and remove something from somewhere before adding something new. In the context of databases, it can be inserted, but depending on how long this section of the lesson becomes, I have to remove or shorten something else, which also has a certain relevance. That's where I see the difficulty"* (Group C - 24.11.23 until 26.11.23, Zeitz).

Ability to integrate (after training): DL & AI: Prerequisites: Continued engagement / preparation / time necessary and getting more knowledge

For unspecified topics, it was stated that one needs more background knowledge than what is documented in the training materials. Furthermore, more engagement, preparation, and work would be necessary in order to integrate topics.

Ability to integrate (after training): DL: Prerequisites: Continued engagement / preparation / time necessary

Integration of DL would be possible after appropriate engagement with the topics.

Ability to integrate (after training): DL: Prerequisites: Continued engagement / preparation / time necessary

A prerequisite for integrating AI is to gain practical experience with AI in order to be able to respond to students' questions about it. More knowledge would also be needed: *"I see AI as a process for myself. Over the last one or two years, I have increasingly understood AI and have gained a certain level of confidence. However, at the moment, I still don't feel really confident enough to teach in the 11th and 12th grades"*

Ability to integrate (after training): AI: Prerequisites: Gaining experience in teaching AI & getting more knowledge, students need to have knowledge about databases first

(Group B - 24.11.23 until 26.11.23, Zeitz). Before you introduce AI into the classroom, students should have knowledge of databases.

Regarding when integration is unclear, it would possible in another case, but this depends on whether teachers have materials available for classroom use. In one federal state, content such as DL and AI is anchored in the framework curriculum and will be integrated by one of the interviewees into the upper secondary level at a later date. In another case, it is currently not possible because certain prerequisites are not met (other content would need to be removed, see prerequisites).

Ability to integrate (after training): Unclear if instantly or not – DL & AI: Possible (e.g., when enough materials or when other content in class is substituted)

Another participant mentioned having the courage to attempt integration, but it's unclear if or when integration could occur. Similarly, the complexity of the topic is addressed – it's seen as a challenge to integrate it into teaching.

Ability to integrate (after training): Unclear if instantly or not – DL & AI: Encouraged to try integration; challenging, because of complexity

Regarding a not immediate integration of DL, a teacher expressed that some exercises from the training could possibly be done next year with an eighth grade class.

Ability to integrate (after training): Not (instantly) capable – DL: Planned at a later time point

A participant may potentially introduce a project on the lifecycle of data. Another project was mentioned by another person: *"I still feel confident in integrating Data Literacy. I may have received some additional insights into certain ideas and projects. For example, I'm currently planning a project and have received some additional input from other colleagues on how to implement it and what other ideas there are. This has helped me because exchanging ideas within the school due to the small subject depart-*

Ability to integrate (after training): Unclear if instantly or not – DL: Project regarding data life cycle, training led to new ideas (e.g., for projects), possible (general statement), Orange3

ment is often not so easy" (Group B - 24.11.23 until 26.11.23, Zeitz). Elsewhere, it is mentioned that one is now familiar with working with the tool Orange3 and can use it in an eleventh-grade mathematics class.

Teaching AI was deemed difficult by another teacher because of a lack of practical experience. Additionally, a lack of knowledge also led to uncertainties regarding teaching AI in the eleventh or twelfth grade.

Ability to integrate (after training): Not (instantly) capable – AI: Not enough experience in teaching or confident / secure enough to teach, only basic / insufficient knowledge through training

But there are also cases where integration of AI is planned for a later time, for example, next year or the year after. It was also mentioned, integration would be possible when students would have knowledge of databases.

Ability to integrate (after training): Not (instantly) capable – AI: Planned at a later time point

Some participants' statements are unclear about when or if they plan integration for AI. For instance, they mention the possibility of constructing a complex related to the topic of AI in the subject of CS, or they provide no specific timelines: *"I would rate my knowledge of AI as rather mediocre, more as a user. Now I have learned how I can use that as a teacher. We have done concrete tasks, some of which are also suitable for my students. I can work on those with them"* (Group C - 24.11.23 until 26.11.23, Zeitz).

Ability to integrate (after training): Unclear if instantly or not – AI: Possible (general statement)

Regarding, unspecified topics, a participant said he felt *"[...] prepared to start teaching for the first time now. Whatever ultimately comes out of it, one must then see afterwards and compare it with practice and then possibly change one's teaching"* (Group B - 24.11.23 to 26.11.23, Zeitz).

Ability to integrate (after training): Unclear if instantly or not – Unspecified: Well prepared, Teaching may need to be adjusted based on initial experiences

Furthermore, it was stated that in the future, the establishment of an elective course is planned. Additionally, for integration into a course, the training was a good first step, even it was not yet sufficient. Further work and preparation would be necessary.

Ability to integrate (after training): Not (instantly) capable – Unspecified: Creation / establishment of an (elective) course; training as good entry point, but only basic / insufficient knowledge through training

Regarding future integration, it was mentioned that it was not yet certain whether the material from the training / lesson plans could really be implemented one to one. Some teachers felt well-prepared for integration without stating if or when they actually plan to integrate the topics; however, it was also acknowledged: *"I believe that confidence only comes when you actually teach it the second or third time. Then you actually see the strengths and weaknesses of your own teaching or your own thoughts. That can then be improved"* (Group B - 24.11.23 until 26.11.23, Zeitz).

Ability to integrate (after training): Unclear if instantly or not – Unspecified: Materials / lessons plans may not be implemented exactly 1:1, well prepared

STEAM Germany 3rd round, length 2.5 days:

In the subject of humanities, AI was used to develop and compare pro and contra positions.

Integration: AI, subject humanities, frequency – more than once: Weighing and illustrating pro and contra positions

Furthermore, AI was used by another teacher to create tasks or coordinate systems for specific assignments. However, the statement does not clarify whether it was solely for lesson planning or if students also worked with it in class. In another case, fake news creation was tested (as images, text, or a combination of both) to reflect on how it works and also to compare; although it was not specified what the comparison was made against.

Integration: AI, subject unclear, frequency – once: Task creation, Produce fake news (as text, image etc.) as trial,

There was a general statement regarding the integration of topics, but it was not specified in any way.

Integration: Unspecified, subject unclear, frequency – unclear: General statement

Using AI, there was the creation *of "[...] an experimental field to produce learning videos, with one group allowed to use support [of AI] and the other group without"*. However, the exact process of this was left open in the focus group interview. The experiment was described as *"amateurish"* (Group A - 26.01.24 until 28.01.24, Meißen), as the teacher realized that he lacked competences. After the training, he would now feel more confident in this content.

Group A:
Integration: AI, subject unclear, frequency – once: Create a learning video with AI (as trial, not in class)
Experiences: Negative: Lack of competences led to feeling of unsureness

In group B, a teacher worked with Fiete[10], an AI-based feedback tool. Although it was emphasized that there were different experiences, only positive aspects were mentioned in the interview. In this context, it was stressed that the students were able to independently articulate what they liked or disliked about the feedback or the tool.

Teacher 1 from group B:
Integration: AI, subject unclear, frequency – unclear: Fiete
Experiences: Students - Positive: Arrived at results / contributed independently

Another teacher from interview group B had texts generated in the classroom using ChatGPT and in the subject of German, *"[...] in the style of a particular author. I compared this with a genuine text by the author, and the students were supposed to work out linguistic features in which the texts differ. **The task afterwards was to formulate a prompt in such a way that a better text would be generated by ChatGPT, closer to the original"*** (Group B - 26.01.24 until 28.01.24, Meißen). However, it turned out that it was hardly possible to achieve (reasonable) results regarding this latter task. The teacher then was uncertain about why this hadn't worked.

Teacher 2 from group B:
Integration: AI, subject language (German), frequency – once: ChatGPT
Experiences: Negative: Hardly came to results

[10] See https://www.fiete.ai (no English version available)

From Group C, a teacher stated having experiences with integration. While it was not explicitly stated what was addressed with the students, *"previous experience has shown that the children are just as baffled and amazed by what is possible. I feel that this amazement surpasses the impulse to engage and understand what is actually happening"* (Group - 26.01.24 until 28.01.24, Meißen).

Group C: Integration: Unspecified, subject unclear, frequency – unclear: General statement Experiences: Students - Neutral: More astonishment than understanding

On the other hand, some teachers expressed that they have not yet integrated topics from DL and/or AI.

No Integration – DL & AI: General statement

Before integration of DL or AI would be feasible, one would need to review the provided material from the training once again. Also, it would be necessary to delve deeper into the topics.

Ability to integrate (after training): DL & AI: Prerequisites: Continued engagement / preperation / time neccessary

Regarding the integration of DL or AI, some interview partners do not feel ready or confident enough to use them in the classroom.

Ability to integrate (after training): Not (instantly) capable – DL & AI: Doesn't feel confident / secure enough to teach

Before integration of AI would be feasible, it would need to be ensured that one can achieve reasonable and reproducible results with tools like ChatGPT, for example, in the case of image generation.

Ability to integrate (after training): AI: Prerequisites: You have to come to results in order to teach (e.g, ChatGPT)

Hence, with reference to the example just mentioned, there are reservations regarding the integration of AI into the classroom.

Ability to integrate (after training): Not (instantly) capable – AI: General statement

Some teachers affirm the question of integrating DL and AI after training and feel confident or more confident about it than before the training. However, it is unclear whether integration would be possible immediately or later. One teacher is eager for integration

Ability to integrate (after training): Unclear if instantly or not – DL & AI: Participants feels confident / safe to use in the classroom, encouraged to try

but assumes that they can only ascertain their own confidence regarding further training topics once the planning phase actually begins.

Some teachers affirm the question of integrating DL and AI after training and feel confident or more confident about it than before the training. However, it is unclear whether integration would be possible immediately or later. One teacher is eager for integration but assumes that they can only ascertain their own confidence regarding further training topics once the planning phase actually begins.

Ability to integrate (after training): Unclear if instantly or not – DL & AI: Grimm´s New Fairy Tales, Abalone

The deconstruction of texts, which was already addressed in the chapter on Teachability and difficulties of teaching DL and AI (Insert chapter number), meaning not generating texts with an AI, but rather scrutinizing AI-generated texts based on certain criteria, is something one interviewee would like to apply in future classes.

Ability to integrate (after training): Unclear if instantly or not – AI: Deconstruct texts with AI instead of creating texts

One participant definitely wants to create teaching materials in a language subject (German). However, it is not clear what content this refers to or what specifically from the training should be utilized, or whether it should be implemented immediately or not. Another teacher confirmed having gained ideas from the training, some of which are likely to be partially implemented.

Ability to integrate (after training): Unclear if instantly or not (Unspecified): Create own material (to (some) of trainings' content + integrate, some of the training content probably will be used

Summary:

In computer science (CS) and STEAM fields, efforts to integrate DL and AI into teaching have revealed a notable gap between training content and classroom application. Educators across various countries have expressed that while training sessions increase their comfort with DL and AI concepts, transitioning this knowledge into effective classroom teaching demands deeper engagement with the material and more

robust support. The general consensus highlights that a single training session, even extending up to seven hours, is inadequate for equipping teachers to effectively teach these advanced topics, especially for those who are new to teaching DL and AI. There's a vocal demand for more detailed course plans and tailored materials that cater specifically to different educational levels to facilitate smoother integration into school curriculums.

STEAM teachers approach DL and AI integration by incorporating these topics across various disciplines, including biology and math models. The training material provided foundational knowledge, but teachers indicated that more detailed and supportive resources are necessary for a more effective and nuanced application in diverse classroom settings. Primary educators face challenges in acquiring in-depth factual and didactic knowledge necessary for teaching complex topics like AI and DL.

In the longer training sessions of 2.5 days, particularly highlighted in the CS Germany data, there are several notable aspects that impact post-training integration of DL and AI into classroom settings:

- practical application and project-based learning: The extended duration allows for comprehensive coverage of practical applications within real-world projects.
- Depth of coverage: longer training sessions offer the opportunity to delve deeper into complex topics. This depth ensures that participants leave with a more nuanced understanding of the subjects, which is crucial for effectively teaching advanced concepts.
- Feedback and iteration: Extended training provides greater opportunities for feedback from trainers and peer interaction, which is vital for refining understanding and teaching strategies.
-

Despite the additional time, there were still challenges noted in fully grasping the technical aspects, such as the complexity of algorithms or the integration of sophisticated AI models like neural networks. Participants expressed a need for ongoing engagement with the material and additional support to effectively integrate these technologies into teaching.

3.5 Steps to bring DL and AI into the classroom

CS in all rounds, length 90 minutes – 7 hours:

In all of the countries, teachers stress urgency to integrate DL and AI into the CS framework curriculum. Some teachers in all of the countries argue to integrate DL and AI into other non-CS subjects as well. Teachers from Germany and Austria, suggested that CS should become a mandatory subject in secondary education. In Lithuania, some suggested starting teaching CS already at the primary level. Educators in all three countries, mentioned the challenges of adjusting current frame-work curriculums to accommodate DL and AI, as the current framework curriculums are already overloaded, hence there is a need to reprioritize existing content. The consensus is that the inclusion of DL and AI in framework curriculums won't instantly equip teachers to teach these subjects proficiently, thereby highlighting the need for teacher training. Potential challenges in incorporating DL and AI into classrooms, particularly related to the framework curriculum, include concerns about overloading it and the risk of not anchoring these topics in it at all. Teachers face obstacles such as a lack of readiness, insufficient knowledge, outdated teaching materials, and pedagogical challenges, with potential deterrent effects due to the complexity of the topics. Resistance to tools like ChatGPT and the need for improved infrastructure add further complexities, like insufficient materials, highlighting the importance of addressing such challenges for effective integration.

STEAM 2-3 rounds, length 1.5-4 hours:

There was support for the careful testing of new content before its inclusion in the framework curriculum, acknowledging legal and technical constraints. The integration of DL and AI into the framework curriculum is approved due to their impact on everyday life (of students) and professional life of teachers and also reflecting technological developments in society. The integration of DL and AI into the curriculum is considered with attention to how this integration is managed, taking into account the existing presence of related topics in media literacy. The interviewees emphasized the necessity of including DL and AI topics in teacher education. The importance of early education in these topics is recognized to equip teachers with essential knowledge for effective teaching and application.

Regarding steps for integrating DL and AI topics into classroom teaching, integration into teacher education and the framework curriculum was considered crucial, with a need for careful consideration of when and how to incorporate these topics. The use of interactive websites/tools, devices, and computer programs for practical and interactive engagement were noted. For this to be feasible, adequate infrastructure needs to be provided. Moreover, it was recommended to offer teachers more time, ready-to-use teaching concepts, to ensure effective integration.

The potential challenges associated with integrating DL and AI into classrooms encompass a lack of infrastructure or the existence of technical issues. Teachers face challenges of high workload and time constraints. The absence of prepared teaching concepts, materials and projects, as well as the need for reprioritisation in the (already saturated) framework curriculum, also pose challenges. Concerns about the slow integration of topics were emphasized.

Primary 3rd round, length 3-4 hours:

There is a general agreement on the inclusion of DL and AI topics in the framework curriculum due to their growing importance. Integration is deemed beneficial as modern technologies, including AI, are integral to everyday life and students

therefore must learn about these topics. There were also some concerns about integrating these topics at the primary school level, suggesting the need for specially trained teachers or experts to help with that. The prerequisites for integrating DL and AI is that teachers also have the knowledge to effectively convey these topics. The integration of DL and AI into teacher education is supported due to the lack of foundational knowledge in primary schools. It would also eliminate the need for subsequent knowledge compensation, ensuring that educators are well-equipped from the start to teach students.

The steps for integration into teaching included advocating for a separate CS subject. Additional measures could involve providing sufficient materials and allowing ample time for teachers. Adopting a gamification approach and using clear examples from daily life to enhance students' understanding of AI in class was deemed beneficial. A gradual integration process is suggested for a smooth and continuous teaching experience.

There was uncertainty about how AI will impact every day and school life. Further potential challenges could include lack of time and preparation of teachers with knowledge to teach the topics, insufficient digital devices / infrastructure, and potential financial constraints for acquiring those and other necessary resources.

CS Germany 1st round, length 2.5 days:

The prerequisite is that the topics can be taught effectively, and therefore, the teachers must possess sufficient knowledge.

Framework curriculum integration: Prerequisites – Knowledge of teachers

The participants advocated for the integration of the topics DL and AI into the framework curriculum. Students should acquire basic knowledge about *"[...] Data Literacy, learning how to deal with data. [...] This should actually be done everywhere. It's not even necessarily about computer science, this is m more like general education"* (Participant E -

Framework curriculum integration: Pro arguments / important / necessary – General agreement, students have to know about DL & AI

25.11.22 until 27.11.22, Heidelberg). Also, students should possess knowledge in AI.

Another participant was unsure whether AI necessarily needs to be anchored in the curriculum. Additionally, aspects of AI are already included in some framework curricula/subjects.

Framework curriculum integration: Neutral – Uncertain if students have to know about AI, DL & AI already integrated in framework curriculum

However, there was also the opinion that framework curriculum integration for AI is not so important: *"[...] I don't really hold much stock in all this curriculum stuff. [...] If you're going to teach it, you have to teach it at a reasonable level, I think. And that includes both the teachers and the students. And forcing it through with a club and then trying to convey a few slogans about AI, that doesn't help anyone"* (Participant E - 25.11.22 until 27.11.22, Heidelberg). DL fits less into the subject of CS.

Framework curriculum integration: Not (so) important – Framework curriculum (integration) of AI overrated / not important, what matters is execution (e.g., quality of teaching), DL doesn't (quite) fit in subject CS

AI should generally be included in the framework curriculum, even in various subjects, but it is also a question of how it would be implemented there. Moreover, curriculum integration also depends on the number of hours allocated.

Framework curriculum integration: Pro or against depending on conditions of framework – it is dependent how AI is implemented, dependent of number of teaching hours

AI should be included in the framework curriculum. DL and AI are connected topics (for DL people should now what AI is).

Framework curriculum integration: Topics to integrate – AI; DL and AI as connected topics

DL should be taught in other subjects besides CS, for example: *"[...] dealing with data itself and this socio-political aspect, that actually belongs, doesn't have to be in computer science, it can be somewhere else. So, I just think there's an incredible amount of stuff that needs to go into computer science in every form and there are incredibly few hours available. And*

Framework curriculum integration: Details for subjects and grade/school levels: DL - Ethical, societal aspects of data/DL in subjects other than CS, DL not really part of CS, but could be part of it, DL should be implemented across different subjects

if you could give any of that to someone else, then that would be it" (Participant D - 25.11.22 until 27.11.22, Heidelberg). But there is also the opinion that working with data can, but does not have to be part of CS. It should be taught in all subjects.

Framework curriculum integration: Details for subjects and grade/school levels: AI - for CS integration, other subjects secondary; upper secondary level, in various subjects

For AI, it was suggested that it would primarily belong to the CS subject and in the upper secondary level, and partially integrated into other subjects. Another opinion is that since CS is not a compulsory subject and the topic of AI is important for all students, it should also be definitely included in other subjects.

Teacher education: Prerequisites - Identifying Interdisciplinary Applications of data/algorithms

As a prerequisite for the integration of topics such as DL and AI, the importance was emphasized, *"[...] to at least become aware of the role algorithms play in our lives, the role data plays in our lives, where it is collected, and what can be done with it"* (Participant A - 25.11.22 until 27.11.22, Heidelberg). These are things that are best taught interdisciplinary; for example, in this context, working with location-based data in the subject of geography could play a role.

Teacher education: Pro arguments / specifics - General agreement, important in order to promote future skills + competencies, integrate as fast as possible, because of (future) relevance of new topics

Generally, there is agreement on integrating DL and AI into teacher education. It is crucial for addressing future essential competencies/themes in the 21st century and must accordingly be conveyed to teachers, such as: *"So both data literacy, meaning data competencies for teachers, but also the ability to foster data competencies in students, should definitely be incorporated into teacher education, and long overdue. Yes, because we've actually been in a data-driven, data-fied, and digitized world for years. So, it's actually absurd that we're only now starting to think about it or that colleagues of mine suddenly say they want to know more about it"* (Participant A - 25.11.22 until 27.11.22, Heidelberg). No more time should be wasted in this regard.

Themes related to DL are considered too difficult: *"So, surveys are something difficult, I think, and they are done way too much. And you would have to teach people [teacher education students] properly beforehand how to conduct reasonable surveys. And we have also noticed that it's not that easy. I don't know what they evaluate in the end, then"* (Participant E - 25.11.22 until 27.11.22, Heidelberg).

Teacher education: Contra arguments: DL / topics of DL (e.g., surveys) too demanding

When it comes to the integration of topics, data competency skills should be incorporated into teacher education to be able to foster them later on with students as well.

Teacher education: Topics: DL should be integrated

Topics such as DL and AI should be situated not only within CS but also within other subjects: *"Naturally, it [CS] forms the basis for understanding what information actually is, what data are, how data are collected, how they are processed, including the practical work with data. But then we quickly enter a realm where I say, actually, topics like AI, when you look at where AI applications exist in various areas of life, and also data competencies are essentially problems that can only be addressed interdisciplinary"* (Participant A - November 25-27, 2022, Heidelberg). The emphasis on subjects varies for another participant; for non-CS-teachers, DL and AI may not play as significant a role as for computer scientists, but at the very least, non-CS teachers should have a rough idea of what AI is.

Teacher education: Subjects: CS and in other subjects besides CS; for Non-CS subject – not as much as for CS, but AI foundation

As part of the integration of topics into the school curriculum, one of the prerequisites is to make the importance of these topics clear to the teachers. It is also necessary that teachers are motivated, as well as them having the necessary fundamental understanding and

School curriculum: Prerequisites: Teachers need to understand the significance of the subject and the capacity to grasp DL & AI, high commitment / motivation of teachers; interdisciplinary integration of DL & AI

capacity to grasp these topics. The topics of DL and AI should be integrated not only in CS but also in many other subjects.

School curriculum: Pro arguments: General agreement

The school curriculum should be adapted*: "And not just when new framework curriculum plans or educational plans come, but I would also mention the KMK [Kultusministerkonferenz – Conference of Ministers of Education] competencies, which are already mandatory. [...]. It's not just about computer science at this point, but it's about all subjects contributing their share to implementing these competencies. And I do see the task of each individual school critically, to actually review their own internal curricula and see what we can sensibly incorporate from them and where we see opportunities to incorporate things like data competency or AI applications from our subject perspective [CS] and/or interdisciplinary exchange in other subjects"* (Participant A - 25.11.22 until 27.11.22, Heidelberg).

School curriculum: Contra arguments: Framework curriculum on which school curriculum is based is vague

The framework curriculum is vague and therefore allows for various interpretations, which is not optimal for solidifying its contents into the school curriculum.

Steps for anchoring in the classroom: Framework curriculum – Search within subjects where data competence is needed; not only foundational knowledge in non-CS-subjects; students don't need to know in-depth details about DL/or AI, knowledge dissemination about DL/AI has to be aligned with subject

Regarding the specific steps for integrating for DL and AI topics into school teaching, one should try to identify application areas in non-CS subjects where it makes sense to collect and analyse data and address it accordingly (for example, it would not be enough to simply interpret a data table). On the other hand, it was also mentioned in the interviews that not every student needs to understand everything in detail, but it should be specifically outlined which competencies within the DL/AI field may be relevant for each subject.

DL and AI should be implemented in teacher education, as otherwise, teachers (both current and prospective) would face difficulties teaching these topics in the school curriculum.

Steps for anchoring in the classroom: Teacher education – General agreement

The competencies for teachers should be clearly defined and monitored: *"A bit like what the KMK strategy envisions, where there are very specific competencies that our students must acquire by the end of their school careers. But we actually need that for teachers as well. [...]. There needs to be an authority that also monitors and enforces this obligation, I believe. [...]. However, I also know that there are federal states [in Germany] where it doesn't matter how much trainings you do. And that is something that is structurally or strategically a real problem. So, there should actually be attention paid to this from a political perspective. As a teacher, you're never fully trained"* (Participant A - 25.11.22 until 27.11.22, Heidelberg). There needs to be broader and meaningful training opportunities, not just for DL and AI, but for the entire spectrum of digital competencies.

Steps for anchoring in the classroom: Trainings – defining/updating competencies, what teachers need to have, central authorities for training (execution, monitoring etc.), create training opportunities (not only on DL/AI, but all digital competencies)

Changes should originate from teachers in a sort of bottom-up process, for which one would need to create a kind of "aha" moment or shock moment for them, for example, by dealing with chatbots via OpenAI (ChatGPT).

Steps for anchoring in the classroom: Other: Bottom-up-process (starting with teachers)

The students should learn basic competencies at the lower secondary level (e.g., interpreting simple Python codes or spreadsheets). Additionally, they should also acquire basic mathematical competencies.

Steps for anchoring in the classroom: Other – Students: Students have to gain computational knowledge and foundational mathematical knowledge

To convey these new topics, teachers need to have the necessary knowledge, for example, through the creation of training opportunities as mentioned above. In addition to subject knowledge, teachers should also be empowered to convey this knowledge to students in a way that is didactically meaningful. Furthermore, educators should bring soft skills such as a willingness for lifelong learning or problem-solving competencies.

Steps for anchoring in the classroom: Other – Teachers: Teachers have to gain knowledge and must be able to teach the subjects, development of soft skills

There should be sufficient material available for use in the classroom. Additionally, materials could be centralised and made easily accessible for all teachers. Access to materials on the internet was also preferred, which would have advantages: *"[...] I don't need a textbook. [...]. There would have to be a concept that is maintained permanently. Some things need to be thrown out, right? That's the case with computer science"* (Participant E - 25.11.22 until 27.11.22, Heidelberg). Furthermore, to be able to teach topics like DL and AI, schools must also have sufficient technological infrastructure.

Steps for anchoring in the classroom: Other – Resources and materials: Sufficient material for teaching, centralise materials, provide websites that are regularly updated, provide infrastructure

When it comes to discussing potential challenges, it would generally be difficult to make people understand what AI is, how it works, and where the areas of risk lie. Additionally, there is a lack of proper data preparation for working with it in schools: *"That is the difficulty of working with data. It would actually fit into many subjects. I could use AI to analyse texts in German just as well. I could model pandemic data in mathematics. Similarly, I could use geographical data or economic data in geography and economics, etc. and so forth. [...]. There are websites, for example, 'World in Data' is a large site that collects data on all sorts of topics and somewhat prepares it, but unfortunately not yet in a form that can be used directly in schools [...]. The various government agencies, the federal government, the states, they have masses of data. Most of the time, it's not*

Potential challenges: Other - Difficult to convey the relevance of AI because it is black box, central database for data for classroom usage missing, institutions could be obstructive

even classified, but actually freely available. There are corresponding data portals where they are available under a Creative Commons license, yet that last step of taking them and using them in schools is still missing" (Participant B - 25.11.22 until 27.11.22, Heidelberg). Furthermore, institutions could be hindering the integration of these topics, but no further details were provided.

If the content taught in teacher education is not sufficiently broad, it is difficult for prospective teachers to conceptualize and teach relevant content for their classes.

Potential challenges: Teacher education – Knowledge is not (broad), so teachers aren't well prepared

Furthermore, the potential obstacles for teachers are related to the acceptance and motivation to teach topics such as DL and AI and their value. If teachers do not possess the necessary knowledge or pedagogical skills to convey these topics, obstacles may arise for them. The implementation of plugged approaches like programming can deter some teachers.

Potential challenges: Teachers – Acceptance of DL/AI by teachers; lack of motivation, knowledge or pedagogical implementability

Another obstacle is the lack of time or high workload, which can result in teachers not being able to manage new (additional) topics that may arise alongside the mandatory curriculum.

Potential challenges: Teachers – High Workload / little time, other priorities

One hurdle is that students do not bring the necessary knowledge: *"So, yes, I experience it again and again, for example, in the ninth grade when students from different classes come. Some already have a bit of knowledge. Some have never heard of it before. You often have to start from scratch and can't say: 'So, now we have a dataset, with, I don't know, 5000 rows and 20 columns, and we'll just throw that into the spreadsheet and do some nice correlation analyses' or something like that, but you actually have*

Potential challenges: Students – Students don't have the computational knowledge, heterogeneous student body regarding knowledge

to start again with these tools. And yes, that's a prerequisite" (Participant B - November 25-27, 2022, Heidelberg).

Potential challenges: Framework curriculum – Decision-makers (for CS) are not subject matter experts, no problems when framework curriculum clearly states that content has to be taught, too less hours for AI in subjects like CS

One obstacle is that those responsible for incorporating CS content into the framework curriculum may not be computer scientists themselves. If the framework curriculum dictates that topics like AI should be taught – and if teachers possess the requisite knowledge (as mentioned above) – then no major issues are anticipated. However, one hindrance may be that in computer science-related subjects, the number of hours allocated may be too limited or insufficient for topics such as AI.

Potential challenges: Time - Lack of time for initiating integration (trainings, testing etc.)

A major issue is generally seen in processes of change within schools, which also applies to the introduction of new topics such as DL and AI: *"First of all, there is no time allocated for that. [...]. It's not just about having a half-day conference with colleagues, these are long-term school and curriculum development processes that need to be initiated. In that sense, what we really need are retreats, where we have several days with colleagues, where we can also try out certain things in teaching, and then come together again to discuss experiences. [...] these are processes that need to be long-term in nature, where feedback is needed in between, where things actually need to be adjusted, discarded, where the school can gain its own experience"* (Participant A - 25.11.22 until 27.11.22, Heidelberg).

Potential challenges: Resources and materials – Lacking materials, devices (like computers) are used in too few subjects

There is a lack of material for the new topics, although there has been a lot of progress in the last two years. Another problem identified regarding resources and materials is that in too few subjects, devices like computers are actually used, and usage is rather limited to subjects that are related to CS.

CS Germany 3rd round, length 2.5 days:

It is noted that framework curriculum integration is linked to the knowledge of teachers. Using the example of mathematics teachers, it is described that they do not utilize spreadsheet calculation in Excel – even though included in the framework curriculum – because they lack the necessary knowledge for it. As long as CS is not a mandatory subject and/or more hours are not dedicated to it, the integration of DL and AI topics into the framework curriculum makes little sense.

Framework curriculum integration: Prerequisites – Knowledge of teachers, CS as mandatory subject or/and with more classroom hours

The integration into the framework curriculum is considered meaningful and necessary, among other reasons, *"[...] because the train has already gone so far in these aspects that we cannot keep pace otherwise"* (Group C - 24.11.23 until 26.11.23, Zeitz). Therefore, the framework curriculum must be brought up to date.

Framework curriculum integration: Pro arguments / important / necessary – General agreement, current framework curriculum needs to be updated / worked on

The integration of topics like AI into the framework curricula is important because such topics have a connection to students' everyday lives, as they already have contact with them, and it also pertains to everyone's daily life in general. For example: *"It is a fundamental everyday skill to be aware that data is collected but never deleted, that it exists, and this begins with children from birth"* (Group A - 24.11.23 until 26.11.23, Zeitz). The framework curriculum therefore should reflect this reality.

Framework curriculum integration: Pro arguments / important / necessary – Students already in contact with DL and AI, topics shapes everyday life, AI is part of everyday / daily life, framework curriculum reflects daily life / new technologies

According to one interview, the integration of DL and AI into the framework curriculum would be unavoidable, but it is not clear how the person feels about it. In another case, DL is said to be already integrated into the framework curriculum and is the second most important topic. Nevertheless, one should consider reorienting the

Framework curriculum integration: Neutral – Integration of topics into framework curriculum is inevitable; although DL already integrated in framework curricula, there could be modifications

topic because currently the compatibility with AI is not fully given. In mathematics, data collection is already heavily emphasized. For the participant, it is unclear whether one should already start working with software in this area at an earlier stage.

Framework curriculum integration: Neutral – Integration of topics into framework curriculum is inevitable; although DL already integrated in framework curricula, there could be modifications

According to one interview, the integration of DL and AI into the framework curriculum would be unavoidable, but it is not clear how the person feels about it. In another case, DL is said to be already integrated into the framework curriculum and is the second most important topic. Nevertheless, one should consider reorienting the topic because currently the compatibility with AI is not fully given. In mathematics, data collection is already heavily emphasized. For the participant, it is unclear whether one should already start working with software in this area at an earlier stage.

Framework curriculum integration: Pro or against depending on conditions of framework - There has to be content reprioritised, curriculum already overloaded

As for integration into the framework curriculum, the question would be about which existing topics would be omitted when adding topics connected to DL or AI, for example: *"We also need to reconsider in which areas we might need to streamline the curricula in general, so that we have the capacity for these additions. We cannot keep piling more material onto our students, as is happening to some extent currently, for instance with the mandatory subject of computer science. We also need to consider the capacity of our students as a resource. How much more can they handle if we keep increasing the workload and tightening the curriculum?"* (Group B - 24.11.23 until 26.11.23, Zeitz).

Framework curriculum integration: Topics to integrate – AI (e.g., Large Language Models like ChatGPT), DL/data competence

Regarding the integration of topics such as AI, as exemplified by ChatGPT (Large Language Model), and data literacy, they would be crucial for all students and their everyday lives. Data literacy should be a global concern and not solely taught within the subject of CS.

When the participants delved into details regarding the integration of DL and/or AI, it was emphasized that these topics should be addressed not only in CS but also in other subjects. Additionally, there should be an adequate number of classroom hours available, for example in CS.

Framework curriculum integration: Details for subjects and grade/school levels: DL or/and AI should be implemented across different subjects, sufficient number of class hours

In regard to the integration of DL and AI into teacher education, only pro-arguments were presented during the group interview, and general approval was expressed. Teachers would need to acquire pertinent knowledge before they can effectively impart it to their students. Additionally, if DL and AI are incorporated into the framework curriculum, these should also be included in teacher education.

Teacher education: Pro arguments / specifics - General agreement; teachers must gain knowledge; in order to teach students, teachers must have knowledge; when in framew. curr., it should be also in teacher education

DL and AI should not only be integrated into CS in teacher education, but also in other subjects. By repeatedly testing practical exercises such as Abalone, teachers could gain experience: *"In doing so, one encounters various problems, solves them somehow, and thereby acquires proficiency in using the tools"* (Group C - 24.11.23 until 26.11.23, Zeitz).

Teacher education: Pro arguments / specifics - AI or/and DL for all teachers, not only CS teachers, practical trial runs (although not exactly clear in which professional phase)

The school management should emphasize that DL or/and AI is incorporated as a cross-cutting theme in the curricula.

School curriculum: DL or/and AI as interdisciplinary topics across subjects

Thus, when considering specific steps for the anchoring of DL and AI in the classroom, integration into the school curriculum was also deemed necessary.

Steps for anchoring in the classroom: School curriculum – Topics have to be in school curriculum

Additionally, the anchoring of DL and AI into the framework curriculum was also considered necessary. DL and AI could be introduced as compulsory or partially compulsory components. They should also be accompanied by sufficient classroom hours dedicated to the subject.

Steps for anchoring in the classroom: Framework curriculum – General agreement, anchoring of DL & AI as (partly) compulsory components, accordingly classroom hours needed for DL & AI

If CS were to become a mandatory subject, as advocated in the interviews, *"[...] then the framework curricula would also shift in general, and [DL] could be more firmly established"* (Group B - 24.11.23 until 26.11.23, Zeitz). However, there is also the opinion that CS does not necessarily need to be established as a compulsory subject, but should be represented across all grade levels. Additionally, it would be desirable that *"[...] this application of AI or the evaluation of data or the interpretation of results, from the data training set and test run, etc., should also be incorporated into other subjects. This is ultimately part of our lives, and I believe that other subjects such as German, religion, or others can and should address the topic"* (Group B - 24.11.23 until 26.11.23, Zeitz).

Steps for anchoring in the classroom: Framework curriculum – CS should be mandatory; CS doesn't need to be mandatory but needs to be in all grade levels; interdisciplinary anchoring of DL & AI

As another step towards establishing DL and AI in school education, the provision of substantial trainings (including for all teachers) was named.

Steps for anchoring in the classroom: Trainings – Create training opportunities

For the integration process, motivated teachers and persons who are responsible for this process, would be needed.

Steps for anchoring in the classroom: Other – Teachers: Teachers / people in charge have to be motivated

There should be sufficient materials (such as textbooks) available to teach DL and AI. Materials should also be centrally accessible or bundled and provided with free access, respectively there should be more Open Educational Resources.

Steps for anchoring in the classroom: Other – Resources and materials: Sufficient material for teaching, centralise materials, more Open Educational Resources materials

Steps for anchoring in the classroom: Other – Resources and materials: User-friendly tools, software and topics; offer browser-based software and software that is installed on PCs

Also important are the user-friendliness of programs, learning tools and the topics taught. Several requirements are mentioned for software, with the first referring to browser-based software that

Interviewee 1: "[...] does not need to be installed. I have a website, ideally free, so that schools and students at home can use it without needing to install anything.

Interviewee 2: And they can use it on different devices. If they have a tablet, they can use it browser-based." (Group A - 24.11.23 until 26.11.23, Zeitz)

On the other hand, software should also be provided, bundled/packaged, and installed on PCs in the classroom.

Potential challenges: Teachers – Acceptance by teachers

For one, the acceptance by the teaching staff could pose problems.

Potential challenges: Students – Students may lack motivation or knowledge for more complex topics (e.g., AI), overwhelming the students

When the focus is on the students, the interviewees also see a risk in the motivation of the students. However, this could depend on the school level; for example, there may be fewer problems in this regard at a gymnasium than at an integrated secondary school (ISS). Therefore, content must be presented in a truly interesting way. Furthermore, students may not have the necessary prior knowledge, such as mathematical backgrounds, or cannot fully grasp topics such as neural networks (AI). Following on from this, there is a risk of overwhelming students. For example, in this context, the topic of linear regression was mentioned: here, one could resort to as simple applications as possible, for instance, using a two-dimensional space instead of a multidimensional one as in the training. Also, technically, through *"[...] software changes and version changes, errors occur*

that quickly overwhelm the students. A lot of frustration arises. One has to invest a lot of time in correcting the errors. The development is so dynamic that it is difficult to implement it in teaching" (Group D - 24.11.23 until 26.11.23, Zeitz).

It remains to be seen how students will accept and deal with the topic of AI: *"One always has to try it first. Then one has to see how many students can cope with certain things. I find this often very difficult, especially in the subject of mathematics. There have been situations where I thought students would find it very difficult, and then they just breezed through it. Other situations, I thought I could do it quickly on the side, and the students had major problems with it. I see the same thing with AI. The shock of practical experience will definitely still be there to some extent"* (Group B - 24.11.23 until 26.11.23, Zeitz). Furthermore, the range of available tools is very large - their relevance for classroom use/students must be tested first. Another problem is finding the right dataset size for the students (dataset too small - not very meaningful; too large - excessive complexity). The datasets would need to have a certain relevance, not only thematically, but also in terms of size, although they must not become too complex, which could be demotivating for the students.

Potential challenges: Students – One must first observe how students cope with new content (AI), abundance of instruments/tools - what is relevant for students?, problems finding right data for students

Further potential obstacles could be rooted in the framework curriculum. In this regard, the differences between the federal states in Germany due to educational federalism were also addressed (such as the starting point of CS in different grade levels). While databases are already part of the framework curriculum, it is unclear to what extent they are already linked to AI; thus, there is uncertainty about the connection of DL to AI, and a possible integration of these two topics via big data is missing in the framework curriculum. Moreover,

Potential challenges: Framework curriculum – General statement, federal differences in education policy/school policy, bridge from AI to DL is missing, framework curriculum full: hardly/no time for other content, CS too less hours

for CS it is already very comprehensive, leaving little time for other topics. For topics like DL and AI, you need a lot of time to get into it. The number of hours in CS is too small to cover the breadth of topics in class.

In general, it takes time to theoretically work through and prepare these topics.

Potential challenges: Time-related – It would take time to engage with and prepare topics

Another obstacle can be that it takes time to install programs like Orange3 on all PCs.

Potential challenges: Administrative – Installation of programs

An infrastructure problem that generally affects the teaching process in CS is that the computers are very slow or outdated.

Potential challenges: Resources and materials – Lack of resources / infrastructure to integrate topics

Possible problems are suspected regarding data protection: *"Which models and applications can I even use with the students? Which ones am I allowed to use at all? That is partly not yet clarified at all"* (Group B - 24.11.23 until 26.11.23, Zeitz). Furthermore, data protection leads to difficulties in finding suitable and relevant datasets for teaching. Tools like Orange3 are too complex, to just use it for individual tasks.

Potential challenges: Other – Data security may prevents usage of specific tools/instruments and finding right datasets for students, complexity of the tools (e.g., Orange3),

Additionally, it is challenging when the school is not oriented towards CS but, for example, has a non-mathematical, natural science focus.

Potential challenges: Other – CS no big relevance in non-mathematical scientific high school

One participant stated that regarding DL, they do not see any major challenges and that integrating this topic into the classroom teaching would therefore be easily manageable.

Potential challenges: Not / hardly expected – For DL no (bigger) challenges seen

STEAM Germany 3rd round, length 2.5 days:

It would be important to find, across the federal states, as unified an understanding as possible of the new topics, also through the supportive influence of research on the ministries of culture, and furthermore, to establish concrete content, like clear competency definitions: *"Looking at these framework curricula, one can see that competency expectations consist of very large word clouds that need to be translated into practice. If I imagine that this takes place in a field that is in flux, that, as we have seen here, is not concretely graspable, then I can already imagine how vague it is formulated for the 16 federal states"* (Group A – 26.01.24 until 28.01.24, Meißen).

Framework curriculum integration: Prerequisites – Standardisations (across federal states), clear definitions (e.g., competences, topics)

Among the interviewed participants, there is agreement on the necessity of integrating DL and AI content into the framework curriculum; however, there are no statements regarding the reasons.

Framework curriculum integration: Pro arguments / important / necessary – General agreement

As content to be integrated, the critical assessment of AI-generated content was mentioned.

Framework curriculum integration: Topics to integrate - (Critical) reflection / risks of AI

The sciences field should thoroughly engage with the topics of DL and AI and develop detailed plans/concepts for series of lessons that can be tested in both training and practice. Prospective and already working teachers could then provide feedback on them.

Teacher education: Prerequisites - Pre-Engagement with topics should be left to experts (scientific field)

Integration into teacher education is also agreed upon. One reason given for this is that teachers need to learn about the content themselves before they can convey it to the students.

Teacher education: Pro arguments / specifics – General agreement; in order to teach students, teachers must have knowledge

Furthermore, while there is agreement for integrating DL and AI into teacher education, equal importance is also emphasized for training programs for already practicing teachers. Additionally, in

Teacher education: Neutral - Teacher education and training equally important; advocated, but responsibilities should not remain with teachers

one interview, there was divided opinion on integration, as it shifts responsibilities towards the teachers: *"[...] but if it's in teacher education, it shifts the focus again. It's good to receive a basic education, but it shouldn't change in a way that expects us to be responsible for it. As an example, I point to the psychological and pedagogical diagnostics that have now been integrated in Berlin. There is simply a lot else that we [teachers] need to cover"* (Group A - 26.01.24 until 28.01.24, Meißen).

Teacher education: Contra arguments: Trainings more important than teacher education

Other opinions do not see a significant importance in teacher education: *"Rather the need to provide ongoing training for existing teachers. I have the impression that people coming from universities often handle it more openly or even bring ideas into schools themselves. Because it's already being used much more intuitively"* (Group C - 26.01.24 until 28.01.24, Meißen). Additionally, too much time would be lost if to start with current teacher trainees, and therefore it would be better to focus on training for in-service teachers who have not yet mastered these topics or have not dealt with them before.

Steps for anchoring in the classroom: Schools and school authorities must inspire enthusiasm, political circles must be contacted

As an important step for anchoring DL and AI in classroom teaching, it is determined that schools and school authorities convey enthusiasm. This could include their participation in information events or trainings on these topics. It is also necessary to engage with political decision-makers to stimulate political will regarding DL and AI and to contribute to their integration into framework curricula.

Steps for anchoring in the classroom: Framework curriculum – DL and AI as mandatory components

When it comes to specific steps for integrating DL and AI, the integration into curriculum frameworks is mentioned again, this time as a compulsory component.

It is advocated to offer (much) more training for in-service teachers.

Steps for anchoring in the classroom: Trainings – Create training opportunities

Among teachers, there should be an establishment of *"a broader knowledge base. By that, I mean that more people need to be informed, especially about the pedagogical basics that must be in place to integrate them [DL and AI]"* (Group B - 26.01.24 until 28.01.24, Meißen). Furthermore, existing concepts regarding assessment mechanisms (regarding exams) for students should be reconsidered. A systemic and financially feasible solution should be created, funded by the respective federal state, for how data protection-compliant work can be carried out in the classroom. Additional resource-related steps include providing technical infrastructure such as functioning Wi-Fi or end devices.

Steps for anchoring in the classroom: Other – Teachers have to gain (pedagogical / didactic) knowledge; students' exams need to thought upon; regarding resources and materials, creating the possibility for data protection-compliant work and provision of infrastructure

In general, the entire establishment of DL and AI should be seen and acted upon as a holistic process: *"Establishing it does not only involve one or two selected colleagues from the school volunteering or being intrinsically motivated to attend training sessions. Instead, it must be carried by everyone. Establishment cannot mean that one person does it while ignoring the others completely"* (Group C - 26.01.24 until 28.01.24, Meißen). This also includes promoting a positive and open approach to such new technologies, which views and treats this process and topics like DL and AI as opportunities rather than just a fear-inducing risk factors.

Steps for anchoring in the classroom: Other – Establishment of DL and AI as holistic process, alleviate people's fears / reservations

Regarding potential challenges, a lack of knowledge among teachers could be a factor: *"I also wonder why teachers are expected to constantly familiarise themselves with all these topics. We are not experts at all in this regard [DL and AI]. We cannot say at all what the societal significance even is. We can inquire about it, we can discuss it"* (Group A - 26.01.24 until 28.01.24, Meißen). The question was also raised regarding the extent to which teachers have an open mind towards these topics. Additionally, a shortage of collegial exchange between schools is already noted, and there is the question of how (potential) exchange between federal states could function.

Potential challenges: Teachers - Lack of knowledge or pedagogical implementability, acceptance by teachers, lack of collegial (possibly institutionalised) exchange

Concerning the students, the heterogeneity of the classes presents challenges.

Potential challenges: Students - Heterogeneous student body regarding knowledge

In relation to the framework curricula, the differences between the federal states due to educational federalism are problematic, as there are no uniform regulations such as common media literacy frameworks: *"As a result, everyone reinvents the wheel in their own school, in their own lessons. In a field that is currently progressing very rapidly and bringing about so much change, it becomes even more difficult to stay up to date, find areas of application, and integrate it meaningfully into teaching"* (Group A - 26.01.24 until 28.01.24, Meißen).

Potential challenges: Framework curriculum - Federal differences in education policy/school policy

In general, it has been said that there may be bureaucratic, administrative or data security related obstacles. It has also been noted that uncertainties may arise in the context of AI due to copyright regulations. When using the free ChatGPT, there may be technical hurdles, as students would need a separate login or the school

Potential challenges: Administrative - Bureaucratic and administrative hurdles (General), copyright issues could arise (in connection to using AI); ChatGPT can't be used (regulation), but alternatives cost money + students would need separate log-in

would need an extra login. Other alternatives such as SchulKI[11] would, in turn, need to be funded.

Potential challenges: Time - In general, lack of time for initiating integration (trainings, testing etc.), too slow integration of topics, rapid pace of topics

An obstacle could be time. For example, integration of topics like DL and AI may not occur if teachers don't have the time available for further trainings. A too slow integration is feared: *"Systematic obstacles from the school system or from above, where people are not yet thinking ahead and still want to teach according to the current curriculum, while they should actually be further ahead. The change will come. One could think ahead, but of course then one is behind again"* (Group B - 26.01.24 until 28.01.24, Meißen). The rapid pace of DL and AI topics will not make it easier to find suitable application areas and to integrate them meaningfully into teaching.

Potential challenges: Resources and materials - Infrastructure not working and technical hurdles, (self-developed) materials cannot be shared

In terms of equipment at schools, it could happen that the infrastructure does not work, for example, if the internet connection fails. Therefore, unplugged materials may also be preferred, as they are technically independent of such malfunctions. Self-developed materials cannot be shared: *"For example, I prepare my upper level students in [subject], make it available to all teachers in [location], and they can modify or improve it as they wish. But that doesn't happen because it is not allowed to provide it or because of a data security issue, because you also use materials that are protected by data security law. […] that's already a barrier"* (Group A - 26.01.24 until 28.01.24, Meißen).

Potential challenges: Other – Schools over self-reliant, systemic barriers in education system or higher authorities, technical hurdles

The schools have to be self-reliant on themselves: *"Everyone […] has to figure it out for themselves, or maybe with two or three others at their school, how to make it work and spread it widely. How do we convey*

[11] See https://schulki.de (no English version available)

this [DL and AI] to our students?" (Group A - 26.01.24 until 28.01.24, Meißen).

Summary:

For CS teachers, the urgency of embedding DL and AI into the CS curriculum is universally acknowledged. There is a strong advocacy for making CS a mandatory subject in secondary education to ensure foundational knowledge is built early. Some educators also support introducing CS at the primary level. The challenges highlighted include an overloaded curriculum, the necessity for teacher training, and infrastructural deficits which hinder the effective integration of new technologies.

STEAM educators recognize the interdisciplinary value of DL and AI, advocating for their integration across various subjects within the STEAM framework. The integration strategy involves testing new content meticulously before its formal inclusion to manage technical and legal challenges effectively. STEAM teachers also highlight the importance of incorporating these topics early in teacher education programs to equip educators with the necessary skills for effective instruction. The challenges for STEAM teachers mirror those of CS educators, focusing on high workloads, insufficient teaching materials, and the need for a more robust infrastructure. Additional challenges include the need for specific applications for specific STEAM subjects.

For primary education, the consensus revolves around the importance of introducing DL and AI concepts at a foundational level. Educators stress the need for specially trained teachers to introduce such advanced topics to young learners. The integration at this level is considered beneficial as it aligns with the technological ubiquity in students' lives. However, the challenge lies in adapting these complex subjects in a manner that is understandable for younger students without overwhelming them. Primary teachers also emphasize the necessity of integrating DL and AI into teacher education to prevent a knowledge gap and ensure that educators are well-prepared from the start. Large classes can be a particular problem for integration DL and AI activities, as students usually need a higher support ratio of teachers.

3.6 Follow-up interviews:

The follow-up interviews conducted with computer science (CS) and STEAM educators and in primary education settings aimed to assess the integration of DL and AI content after training sessions. Here's a summary of the findings across different countries and disciplines:

CS:

Germany: Mixed outcomes were reported. One teacher in Berlin seemed to have only implicitly integrated DL into database courses, while AI content was not integrated due to the analog nature of training exercises and the perception that students prefer more hands-on computer work. Another participant noted a lack of time in the curriculum as a major barrier to integrating these topics.

Lithuania: A teacher successfully integrated AI through gamification techniques in teaching fourth graders, demonstrating adaptability in applying training methodologies.

Austria: A participant saw the training as a confirmation to resume previously taught content like machine learning using Orange3, and planned to expand on topics like reinforcement learning. Discussions on integrating AI into ethics through media ethics were also noted.

STEAM:

Austria: In university settings, AI was successfully integrated into teacher training and educational practices. However, Orange3 was perceived as too complex and not further used outside of the training context.

Primary Education

Lithuania: Technology projects involving AI apps for nature exploration showed successful integration of AI into the curriculum.

Austria: The integration focused on imparting basic computer literacy and critical thinking about information sources. Challenges were noted in integrating more complex topics due to language barriers among non-native speakers.

In summary, the integration of DL and AI post-training shows varied results across different educational settings. While some educators have successfully adapted and applied the training content, others face challenges like curriculum constraints, the complexity of tools, and the need for more practical, hands-on applications to engage students effectively.

4. Discussion of key findings across the trainings

The data from both quantitative and qualitative analyses suggest that the DL and AI trainings were highly appreciated by the trained teachers across all the three target groups: CS, STEAM, and primary teachers.

Data collected from pre-, post-, and follow-up surveys indicated several key outcomes:

Competence in Teaching DL/AI: There was a notable increase in teachers' self-reported competence in teaching DL and AI content following the training. This was particularly pronounced among Computer Science (CS) teachers for AI topic.

Understanding of DL/AI Concepts: Improvements were observed in teachers' understanding of DL and AI concepts post-training. Despite the complexity of some topics, such as regression and using tools like Orange3, the overall trend showed enhanced comprehension across all target groups.

Attitudes and Motivation: The societal importance of DL and AI was recognized highly by all participant groups. The motivation to further learn about DL and AI remained high post-training, indicating a sustained interest in these fields.

Integration of DL/AI Post-Training: The follow-up data, although limited by low response rates, suggested a positive trend towards the integration of DL and AI into teaching practices, with a number of teachers reporting actual integration of these topics into their classrooms.

Qualitative insights were garnered from interviews and focus groups post-training. These responses highlighted several important themes:

Training Effectiveness: Teachers appreciated the practical applications and the alignment of the training with their existing knowledge levels. However, they noted the need for deeper and more comprehensive training sessions to fully grasp complex DL and AI concepts.

Barriers to Integration: While the training was generally well-received, some teachers expressed challenges in integrating the learned content into their teaching practices. The need for more tailored content, especially for STEAM and primary teachers, was evident, as they need more concrete applications for their subjects.

Suggestions for Improvement: Participants suggested several improvements for future trainings, including extending the duration of the sessions using various formats, incorporating more hands-on exercises, and providing more detailed content that is directly applicable to various teaching settings.

Both quantitative and qualitative findings suggest that the 2.5-day training sessions allowed for a more thorough exploration of complex topics, significantly enhancing participant engagement and understanding through varied instructional strategies and practical exercises. While the extended duration facilitated deeper learning and provided ample opportunity for reflection and feedback, challenges such as participant fatigue, logistical constraints, and the intense focus on certain content areas were noted. Feedback suggested the need for balanced scheduling to manage cognitive load, customization of content to cater to varying expertise levels,

and follow-up support to reinforce learning and application in teaching practices. Overall, these 2.5 trainings underscored the benefits of longer training durations while highlighting areas for refinement to maximize effectiveness and accessibility.

The high recognition of the societal importance of DL and AI and the motivation to learn more about these fields among teachers suggest a robust foundation for the broader integration of these topics into educational systems.

5. Limitations

Some inherent limitations of the evaluation should be considered:

- Low registration numbers for primary/STEAM trainings: there registration numbers for Primary and STEAM trainings were comparatively low. Lithuanian team reached out to schools for the primary trainings, which resulted in a relatively high number of participants for the primary training.

- Self-selection of teachers: teachers who signed up for the training might differ from teachers who did want/could not participate.

- Follow-up response rates: contacting participants for a 6-month follow-up to evaluate teaching DL/AI in class was challenging. The sample sizes for the follow-up-survey (with an exception of CS and Primary in Lithuania) were very low.

- Training format variation: While longer trainings (2.5) were added to look into the possibility of longer trainings, they were conducted only in Germany and only for CS and STEAM. Other real-world constraints mentioned in other deliverables and the reporting made it hard to organize the format of the trainings so that a systematic variation in the format, topic, location, and the target

group could be controlled for. Hence, the evaluation design was changed to exploration, where each format and training was analysed more in depth.

- Sample size: The small number of participants from each location affects the generalizability of the findings. Expanding the sample size in future studies would contribute to robust and generalisable outcomes.

- Limited sample: various local contexts and school types within the countries were not sufficiently covered.

- Validity and reliability of data collection instruments: Due to time and re-courses allocation, the quality of the survey and knowledge test instruments was not evaluated prior to their application, which can influence the reliability of the data collected.

- Cross-Cultural Challenges: While comparing findings between the countries, one should be aware of the fact that differences could also be traced to cross-cultural differences affecting survey response styles, e.g., differences in acqui-escence re-sponse styles (Rammstedt, Danner, and Bosnjak 2017).

- Differentiation between subjects: Due to small sample sizes, it was not possible to make statements for certain subjects (apart from differentiating CS from STEAM an primary subjects), yet there might be relevant differences in how DL and AI can be integrated in various STEAM subjects

6. Conclusion and future questions

The TrainDL project's three intervention rounds have effectively contributed to the strategic goal of developing policy recommendations for integrating DL and AI into teacher trainings. This final report synthesizes outcomes and challenges based on the implemented trainings for the three target groups: CS, STEAM, and primary in-service teachers, including the longer 2.5 day trainings.

Key achievements:

The TrainDL project's third intervention round has effectively contributed to the strategic goal of developing policy recommendations for integrating DL and AI into teacher trainings. This final evaluation synthesizes outcomes and challenges based on the implemented trainings for the three target groups: CS, STEAM, and primary in-service teachers.

Broad Reach: The project successfully reached out to all of the three target groups (CS, STEAM, and primary teachers) as planned in the project proposal.

High participant satisfaction: The overwhelming majority of trained teachers were satisfied with the training and provided overwhelmingly positive feedback.

Effective Training Materials and Exercises: The training exercises were particularly well-received, with most participants rating the introduced activities highly. This indicates that most selected activities were deemed highly suitable and relevant for teaching. This positive response to the training exercises demonstrates their practical applicability and relevance for teaching for all of the three target groups.

Increased Confidence in DL and AI Application: Participants across all of the target groups reported that the TrainDL training served as a valuable introduction to AI and DL. Despite a higher initial familiarity with AI over DL, the training effectively improved

participants' confidence in applying both topics in their teaching practices, suggesting that the sessions effectively addressed some of the foundational gaps in DL and AI knowledge.

Recognition of educational and educational value of DL and AI across all the target groups: Despite varying initial familiarity levels with DL and AI, there was a notable enthusiasm among participants in all of the three target groups to learn further as well as deepen their knowledge, indicating a successful recognition of the importance of these technologies in education. Participants showed a particular interest in integrating AI into the curriculum as well as their teaching, although there was a noted underappreciation of DL's foundational role in understanding AI.

Recognition of value of (further) teacher trainings: Teachers emphasize the need to incorporate DL and AI into the CS framework curriculum, yet merely adding DL and AI to the curriculums won't ensure that teachers are immediately adept at teaching them, also underscoring the need for teacher training.

Key Challenges:

Finding suitable format and length: single trainings are not sufficient. The 2.5 training provided more opportunities for more in-depth engagement with the material, including project-based learning, feedback from the training and interaction with the peers. Yet, even participants of the 2.5 trainings expressed a strong demand for continuing their training. To address the challenge of teachers' limited time, future trainings could be designed as modular and (where possible) blended learning trainings and spread over multiple sessions. Development of specific course plans and materials tailored for different grade levels is essential. This step will help streamline the integration of DL and AI into teaching, reducing the preparation workload for teachers.

Lacking technical infrastructure and resources: Technical infrastructure sufficient for DL and AI education is lacking in some schools. It is essential to provide stable and sufficient technical infrastructure as well as technical support.

Challenges for integration into framework curriculum: Framework curriculum integration was supported, but is dependent on addressing other challenges such as overloaded framework curricula, ensuring feasibility of new content sufficient trainings of teachers.

(only STEAM and Primary) No clear connection of DL and AI to the subjects: teachers did not always see a clear connection of DL and AI to their subjects and have to prioritize their curriculum content, leaving no space and no incentives to integrate DL and AI subjects into their classes. Further teacher trainings need to focus on developing and teaching DL and AI applications and materials for concrete STEAM and primary subjects taking into account appropriate (for the level of students) instructional strategies.

Insufficient knowledge and understanding of the role of DL: The data revealed differences in the initial knowledge and attitudes towards DL compared to AI, with a general tendency to undervalue DL. This suggests a need for more foundational trainings on DL as well as linkage between DL and AI. The 2.5 day trainings for CS managed to convey the role of DL better than the shorter trainings.

Future questions and approaches for evaluations:

Longitudinal perspective: longitudinal studies to track the long-term impact of DL/AI trainings on teaching DL/AI in schools and students' skills perceptions of AI could provide deep insights into the question of integration of DL and AI into teaching.

Multiple perspectives: studies that cover not only perspective of and impact on teachers, but also school administration and school students, could provide a more complete picture.

Beyond self-selection bias: it would be highly beneficial to design studies to cover not only those teachers, who sign up for trainings voluntarily (they are likely to have higher motivation and prior knowledge), but to design studies with selection of random samples or exploration of less motivated teachers

Investigating incentives and structural anchoring of teacher trainings given high workload: As it is very important to provide training formats suitable the needs and constraints of the teachers, it could be very fruitful to explore the question of incentives and best practices for institutionally integrating teacher training.

Focus on pre-service: Due to the feedback of the stakeholders and limited capacities, the focus of the TrainDL was on in-service teachers with only one evaluated pre-service training, which was evaluated only with a post-evaluation survey. Extending and focusing the sample on pre-service teachers can be very valuable to have a complete picture of teacher training.

More applications for STEAM/Primary: There is a need for more evaluation on the topic of how DL and AI can be integrated into primary and STEAM subjects: best-practices (chemistry, physics, primary subjects with the respective pedagogical approaches) and concrete application.

Disclaimer: Parts of this text could be generated or rephrased by ChatGPT, DeepL Write, LanguageTool, and Google Docs spell checking, but were carefully checked and revised by the authors.

7. References

Baskerville, Richard L., and A. Trevor Wood-Harper. 1996. “A Critical Perspective on Action Research as a Method for Information Systems Research.” *J Inf Technol* 11 (3): 235–46. https://doi.org/10.1080/026839696345289.

Burns, Anne. 2010. *Doing Action Research in English Language Teaching: A Guide for Practitioners*. ESL & applied linguistics professional series. New York: Routledge.

Creswell, John W., and Vicki L. Plano Clark. 2018. *Designing and Conducting Mixed Methods Research*. Third edition. Thousand Oaks, California: SAGE.

European Commission / EACEA / Eurydice. “Informatics Education at School in Europe.” https://redined.educacion.gob.es/xmlui/bitstream/handle/11162/235704/informatics.pdf?sequence=1. Accessed June 16, 2023.

Kuckartz, Udo, and Stefan Rädiker. 2020. *Fokussierte Interviewanalyse mit MAXQDA: Schritt für Schritt*. Springer eBook Collection. Wiesbaden: Springer VS.

Mayring, Philipp. 2010. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 11., aktual. und überarb. Aufl. Pädagogik. Weinheim, Basel: Beltz.

OECD. 2023a. “Education GPS - Germany - Overview of the Education System (EAG 2022).” Accessed June 10, 2023. https://gpseducation.oecd.org/CountryProfile?plotter=h5&primaryCountry=DEU&treshold=5&topic=EO.

OECD. 2023b. “Education GPS - Lithuania - Overview of the Education System (EAG 2022).” Accessed June 16, 2023. https://gpseducation.oecd.org/CountryProfile?plotter=h5&primaryCountry=LTU&treshold=5&topic=EO.

OECD. 2023c. “Education GPS - Austria - Overview of the Education System (EAG 2022).” Accessed June 19, 2023. https://gpseducation.oecd.org/CountryProfile?plotter=h5&primaryCountry=AUT&treshold=5&topic=EO.

Rammstedt, Beatrice, Daniel Danner, and Michael Bosnjak. 2017. “Acquiescence Response Styles: A Multilevel Model Explaining Individual-Level and Country-Level Differences.” *Personality and Individual Differences* 107:190–94. https://doi.org/10.1016/j.paid.2016.11.038.

Ridsdale, Chantel, James Rothwell, Mike Smit, Michael Bliemel, Dean Irvine, Dan Kelley, Stan Matwin, Brad Wuetherick, and Hossam ?Ali-Hassan. 2015. “Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report.” Accessed July 16, 2023. https://www.researchgate.net/publication/284029915_Strategies_and_Best_Practices_for_Data_Literacy_Education_Knowledge_Synthesis_Report.

Schröder, Eike, Felix Suessenbach, and Mathias Winde. 2022. “Policy Paper: Informatikunterricht: Lückenhaft Und Unterbesetzt: Informatikunterricht in Deutschland – Ein Flickenteppich Auch Hinsichtlich Der Datenlage.” Accessed June 07, 2023. https://www.stifterverband.org/medien/informatikunterricht.

Schwarz, Richard, Lutz Hellmig, and Steffen Friedrich. 2022. “Informatik-Monitor.” https://informatik-monitor.de/fileadmin/GI/Projekte/Informatik-Monitor/Informatik-Monitor_2022/Informatik-Monitor_2022_2023_FINAL.pdf. Accessed June 07, 2023.